(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT CO

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/44321 A2

(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/EP01/13968

(22) International Filing Date:
29 November 2001 (29.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

00126325.0	1 December 2000 (01.12.2000)	EP
60/279,661	30 March 2001 (30.03.2001)	US
PCT/US01/10188	30 March 2001 (30.03.2001)	US

(71) Applicant *(for all designated States except US)*: **MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V.** [DE/DE]; Hofgartenstrasse 2, 80539 München (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TUSCHL, Thomas** [DE/DE]; Keplerstrasse 9, 37085 Göttingen (DE). **EL-BASHIR, Sayda** [DE/DE]; Sultebecksbreite 2, 37075 Göttingen (DE). **LENDECKEL, Winfried** [DE/DE]; Blinde Gasse 52, 37318 Hohengandern (DE).

(74) Agent: **WEICKMANN & WEICKMANN**; Postfach 860 820, 81635 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/44321 A2

(54) Title: RNA INTERFERENCE MEDIATING SMALL RNA MOLECULES

(57) Abstract: Double-stranded RNA (dsRNA) induces sequence-specific post-transcriptional gene silencing in many organisms by a process known as RNA interference (RNAi). Using a Drosophila in vitro system, we demonstrate that 19-23 nt short RNA fragments are the sequence-specific mediators of RNAi. The short interfering RNAs (siRNAs) are generated by an RNase III-like processing reaction from long dsRNA. Chemically synthesized siRNA duplexes with overhanging 3' ends mediate efficient target RNA cleavage in the lysate, and the cleavage site is located near the center of the region spanned by the guiding siRNA. Furthermore, we provide evidence that the direction of dsRNA processing determines whether sense or antisense target RNA can be cleaved by the produced siRNP complex.

RNA Interference Mediating Small RNA molecules

Description

The present invention relates to sequence and structural features of double-stranded (ds)RNA molecules required to mediate target-specific nucleic acid modifications such as RNA-interference and/or DNA methylation.

The term "RNA interference" (RNAi) was coined after the discovery that injection of dsRNA into the nematode C. elegans leads to specific silencing of genes highly homologous in sequence to the delivered dsRNA (Fire et al., 1998). RNAi was subsequently also observed in insects, frogs (Oelgeschlager et al., 2000), and other animals including mice (Svoboda et al., 2000; Wianny and Zernicka-Goetz, 2000) and is likely to also exist in human. RNAi is closely linked to the post-transcriptional gene-silencing (PTGS) mechanism of co-suppression in plants and quelling in fungi (Catalanotto et al., 2000; Cogoni and Macino, 1999; Dalmay et al., 2000; Ketting and Plasterk, 2000; Mourrain et al., 2000; Smardon et al., 2000) and some components of the RNAi machinery are also necessary for post-transcriptional silencing by co-suppression (Catalanotto et al., 2000; Dernburg et al., 2000; Ketting and Plasterk, 2000). The topic has also been reviewed recently (Bass, 2000; Bosher and Labouesse, 2000; Fire, 1999; Plasterk and Ketting, 2000; Sharp, 1999; Sijen and Kooter, 2000), see also the entire issue of Plant Molecular Biology, vol. 43, issue 2/3, (2000).

In plants, in addition to PTGS, introduced transgenes can also lead to transcriptional gene silencing via RNA-directed DNA methylation of cytosines (see references in Wassenegger, 2000). Genomic targets as short as 30 bp are methylated in plants in an RNA-directed manner (Pelissier, 2000). DNA methylation is also present in mammals.

The natural function of RNAi and co-suppression appears to be protection of the genome against invasion by mobile genetic elements such as retro-transposons and viruses which produce aberrant RNA or dsRNA in the host cell when they become active (Jensen et al, 1999; Ketting et al., 1999; Ratcliff et al., 1999; Tabara et al., 1999). Specific mRNA degradation prevents transposon and virus replication although some viruses are able to overcome or prevent this process by expressing proteins that suppress PTGS (Lucy et al., 2000; Voinnet et al., 2000).

DsRNA triggers the specific degradation of homologous RNAs only within the region of identity with the dsRNA (Zamore et al., 2000). The dsRNA is processed to 21-23 nt RNA fragments and the target RNA cleavage sites are regularly spaced 21-23 nt apart. It has therefore been suggested that the 21-23 nt fragments are the guide RNAs for target recognition (Zamore et al., 2000). These short RNAs were also detected in extracts prepared from D. melanogaster Schneider 2 cells which were transfected with dsRNA prior to cell lysis (Hammond et al., 2000), however, the fractions that displayed sequence-specific nuclease activity also contained a large fraction of residual dsRNA. The role of the 21-23 nt fragments in guiding mRNA cleavage is further supported by the observation that 21-23 nt fragments isolated from processed dsRNA are able, to some extent, to mediate specific mRNA degradation (Zamore et al., 2000). RNA molecules of similar size also accumulate in plant tissue that exhibits PTGS (Hamilton and Baulcombe, 1999).

Here, we use the established Drosophila in vitro system (Tuschl et al., 1999; Zamore et al., 2000) to further explore the mechanism of RNAi. We demonstrate that short 21 and 22 nt RNAs, when base-paired with 3' overhanging ends, act as the guide RNAs for sequence-specific mRNA degradation. Short 30 bp dsRNAs are unable to mediate RNAi in this system because they are no longer processed to 21 and 22 nt RNAs. Furthermore, we defined the target RNA cleavage sites relative to the 21 and

22 nt short interfering RNAs (siRNAs) and provide evidence that the direction of dsRNA processing determines whether a sense or an antisense target RNA can be cleaved by the produced siRNP endonuclease complex. Further, the siRNAs may also be important tools for transcriptional modulating, e.g. silencing of mammalian genes by guiding DNA methylation.

Further experiments in human in vivo cell culture systems (HeLa cells) show that double-stranded RNA molecules having a length of preferably from 19-25 nucleotides have RNAi activity. Thus, in contrast to the results from Drosophila also 24 and 25 nt long double-stranded RNA molecules are efficient for RNAi.

The object underlying the present invention is to provide novel agents capable of mediating target-specific RNA interference or other target-specific nucleic acid modifications such as DNA methylation, said agents having an improved efficacy and safety compared to prior art agents.

The solution of this problem is provided by an isolated double-stranded RNA molecule, wherein each RNA strand has a length from 19-25, particularly from 19-23 nucleotides, wherein said RNA molecule is capable of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation. Preferably at least one strand has a 3'-overhang from 1-5 nucleotides, more preferably from 1-3 nucleotides and most preferably 2 nucleotides. The other strand may be blunt-ended or has up to 6 nucleotides 3' overhang. Also, if both strands of the dsRNA are exactly 21 or 22 nt, it is possible to observe some RNA interference when both ends are blunt (0 nt overhang). The RNA molecule is preferably a synthetic RNA molecule which is substantially free from contaminants occurring in cell extracts, e.g. from Drosophila embryos. Further, the RNA molecule is preferably substantially free from any non-target-specific contaminants, particularly non-target-specific RNA molecules e.g. from contaminants occuring in cell extracts.

Further, the invention relates to the use of isolated double-stranded RNA molecules, wherein each RNA strand has a length from 19-25 nucleotides, for mediating, target-specific nucleic acid modifications, particularly RNAi, in mammalian cells, particularly in human cells.

Surprisingly, it was found that synthetic short double-stranded RNA molecules particularly with overhanging 3'-ends are sequence-specific mediators of RNAi and mediate efficient target-RNA cleavage, wherein the cleavage site is located near the center of the region spanned by the guiding short RNA.

Preferably, each strand of the RNA molecule has a length from 20-22 nucleotides (or 20-25 nucleotides in mammalian cells), wherein the length of each strand may be the same or different. Preferably, the length of the 3'-overhang reaches from 1-3 nucleotides, wherein the length of the overhang may be the same or different for each strand. The RNA-strands preferably have 3'-hydroxyl groups. The 5'-terminus preferably comprises a phosphate, diphosphate, triphosphate or hydroxyl group. The most effective dsRNAs are composed of two 21 nt strands which are paired such that 1-3, particularly 2 nt 3' overhangs are present on both ends of the dsRNA.

The target RNA cleavage reaction guided by siRNAs is highly sequence-specific. However, not all positions of a siRNA contribute equally to target recognition. Mismatches in the center of the siRNA duplex are most critical and essentially abolish target RNA cleavage. In contrast, the 3' nucleotide of the siRNA strand (e.g. position 21) that is complementary to the single-stranded target RNA, does not contribute to specificity of the target recognition. Further, the sequence of the unpaired 2-nt 3' overhang of the siRNA strand with the same polarity as the target RNA is not critical for target RNA cleavage as only the antisense siRNA strand guides target recognition. Thus, from the single-stranded overhanging nucleotides only the

penultimate position of the antisense siRNA (e.g. position 20) needs to match the targeted sense mRNA.

Surprisingly, the double-stranded RNA molecules of the present invention exhibit a high in vivo stability in serum or in growth medium for cell cultures. In order to further enhance the stability, the 3'-overhangs may be stablized against degradation, e.g. they may be selected such that they consist of purine nucleotides, particularly adenosine or guanosine nucleotides. Alternatively, substitution of pyrimidine nucleotides by modified analogues, e.g. substitution of uridine 2 nt 3' overhangs by 2'-deoxythymidine is tolerated and does not affect the efficiency of RNA interference. The absence of a 2' hydroxyl significantly enhances the nuclease resistance of the overhang in tissue culture medium.

In an especially preferred embodiment of the present invention the RNA molecule may contain at least one modified nucleotide analogue. The nucleotide analogues may be located at positions where the target-specific activity, e.g. the RNAi mediating activity is not substantially effected, e.g. in a region at the 5'-end and/or the 3'-end of the double-stranded RNA molecule. Particularly, the overhangs may be stabilized by incorporating modified nucleotide analogues.

Preferred nucleotide analogues are selected from sugar- or backbone-modified ribonucleotides. It should be noted, however, that also nucleobase-modified ribonucleotides, i.e. ribonucleotides, containing a non-naturally occurring nucleobase instead of a naturally occurring nucleobase such as uridines or cytidines modified at the 5-position, e.g. 5-(2-amino)propyl uridine, 5-bromo uridine; adenosines and guanosines modified at the 8-position, e.g. 8-bromo guanosine; deaza nucleotides, e.g. 7-deaza-adenosine; O- and N-alkylated nucleotides, e.g. N6-methyl adenosine are suitable. In preferred sugar-modified ribonucleotides the 2' OH-group is repla-

ced by a group selected from H, OR, R, halo, SH, SR, NH_2, NHR, NR_2 or CN, wherein R is C_1-C_6 alkyl, alkenyl or alkynyl and halo is F, Cl, Br or I. In preferred backbone-modified ribonucleotides the phosphoester group connecting to adjacent ribonucleotides is replaced by a modified group, e.g. of phosphothioate group. It should be noted that the above modifications may be combined.

The sequence of the double-stranded RNA molecule of the present invention has to have a sufficient identity to a nucleic acid target molecule in order to mediate target-specific RNAi and/or DNA methylation. Preferably, the sequence has an identity of at least 50%, particularly of at least 70% to the desired target molecule in the double-stranded portion of the RNA molecule. More preferably, the identity is at least 85% and most preferably 100% in the double-stranded portion of the RNA molecule. The identity of a double-stranded RNA molecule to a predetermined nucleic acid target molecule, e.g. an mRNA target molecule may be determined as follows:

$$I = \frac{n}{L} \times 100$$

wherein I is the identity in percent, n is the number of identical nucleotides in the double-stranded portion of the ds RNA and the target and L is the length of the sequence overlap of the double-stranded portion of the dsRNA and the target.

Alternatively, the identity of the double-stranded RNA molecule to the target sequence may also be defined including the 3' overhang, particularly an overhang having a length from 1-3 nucleotides. In this case the sequence identity is preferably at least 50%, more preferably at least 70% and most preferably at least 85% to the target sequence. For example, the nucleotides from the 3' overhang and up to 2 nucleotides from the 5' and/or 3' terminus of the double strand may be modified without significant loss of activity.

The double-stranded RNA molecule of the invention may be prepared by a method comprising the steps:

(a) synthesizing two RNA strands each having a length from 19-25, e.g. from 19-23 nucleotides, wherein said RNA strands are capable of forming a double-stranded RNA molecule, wherein preferably at least one strand has a 3'-overhang from 1-5 nucleotides,

(b) combining the synthesized RNA strands under conditions, wherein a double-stranded RNA molecule is formed, which is capable of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation.

Methods of synthesizing RNA molecules are known in the art. In this context, it is particularly referred to chemical synthesis methods as described in Verma and Eckstein (1998).

The single-stranded RNAs can also be prepared by enzymatic transcription from synthetic DNA templates or from DNA plasmids isolated from recombinant bacteria. Typically, phage RNA polymerases are used such as T7, T3 or SP6 RNA polymerase (Milligan and Uhlenbeck (1989)).

A further aspect of the present invention relates to a method of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation in a cell or an organism comprising the steps:

(a) contacting the cell or organism with the double-stranded RNA molecule of the invention under conditions wherein target-specific nucleic acid modifications may occur and

(b) mediating a target-specific nucleic acid modificiation effected by the double-stranded RNA towards a target nucleic acid having a

sequence portion substantially corresponding to the double-stranded RNA.

Preferably the contacting step (a) comprises introducing the double-stranded RNA molecule into a target cell, e.g. an isolated target cell, e.g. in cell culture, a unicellular microorganism or a target cell or a plurality of target cells within a multicellular organism. More preferably, the introducing step comprises a carrier-mediated delivery, e.g. by liposomal carriers or by injection.

The method of the invention may be used for determining the function of a gene in a cell or an organism or even for modulating the function of a gene in a cell or an organism, being capable of mediating RNA interference. The cell is preferably a eukaryotic cell or a cell line, e.g. a plant cell or an animal cell, such as a mammalian cell, e.g. an embryonic cell, a pluripotent stem cell, a tumor cell, e.g. a teratocarcinoma cell or a virus-infected cell. The organism is preferably a eukaryotic organism, e.g. a plant or an animal, such as a mammal, particularly a human.

The target gene to which the RNA molecule of the invention is directed may be associated with a pathological condition. For example, the gene may be a pathogen-associated gene, e.g. a viral gene, a tumor-associated gene or an autoimmune disease-associated gene. The target gene may also be a heterologous gene expressed in a recombinant cell or a genetically altered organism. By determinating or modulating, particularly, inhibiting the function of such a gene valuable information and therapeutic benefits in the agricultural field or in the medicine or veterinary medicine field may be obtained.

The dsRNA is usually administered as a pharmaceutical composition. The administration may be carried out by known methods, wherein a nucleic acid is introduced into a desired target cell in vitro or in vivo. Commonly

used gene transfer techniques include calcium phosphate, DEAE-dextran, electroporation and microinjection and viral methods (Graham, F.L. and van der Eb, A.J. (1973) Virol. 52, 456; McCutchan, J.H. and Pagano, J.S. (1968), J. Natl. Cancer Inst. 41, 351; Chu, G. et al (1987), Nucl. Acids Res. 15, 1311; Fraley, R. et al. (1980), J. Biol. Chem. 255, 10431; Capecchi, M.R. (1980), Cell 22, 479). A recent addition to this arsenal of techniques for the introduction of DNA into cells is the use of cationic liposomes (Felgner, P.L. et al. (1987), Proc. Natl. Acad. Sci USA 84, 7413). Commercially available cationic lipid formulations are e.g. Tfx 50 (Promega) or Lipofectamin2000 (Life Technologies).

Thus, the invention also relates to a pharmaceutical composition containing as an active agent at least one double-stranded RNA molecule as described above and a pharmaceutical carrier. The composition may be used for diagnostic and for therapeutic applications in human medicine or in veterinary medicine.

For diagnostic or therapeutic applications, the composition may be in form of a solution, e.g. an injectable solution, a cream, ointment, tablet, suspension or the like. The composition may be administered in any suitable way, e.g. by injection, by oral, topical, nasal, rectal application etc. The carrier may be any suitable pharmaceutical carrier. Preferably, a carrier is used, which is capable of increasing the efficacy of the RNA molecules to enter the target-cells. Suitable examples of such carriers are liposomes, particularly cationic liposomes. A further preferred administration method is injection.

A further preferred application of the RNAi method is a functional analysis of eukaryotic cells, or eukaryotic non-human organisms, preferably mammalian cells or organisms and most preferably human cells, e.g. cell lines such as HeLa or 293 or rodents, e.g. rats and mice. By transfection with suitable double-stranded RNA molecules which are homologous to a prede-

termined target gene or DNA molecules encoding a suitable double-stranded RNA molecule a specific knockout phenotype can be obtained in a target cell, e.g. in cell culture or in a target organism. Surprisingly it was found that the presence of short double-stranded RNA molecules does not result in an interferon response from the host cell or host organism.

Thus, a further subject matter of the invention is a eukaryotic cell or a eukaryotic non-human organism exhibiting a target gene-specific knockout phenotype comprising an at least partially deficient expression of at least one endogeneous target gene wherein said cell or organism is transfected with at least one double-stranded RNA molecule capable of inhibiting the expression of at least one endogeneous target gene or with a DNA encoding at least one double stranded RNA molecule capable of inhibiting the expression of at least one endogeneous target gene. It should be noted that the present invention allows a target-specific knockout of several different endogeneous genes due to the specificity of RNAi.

Gene-specific knockout phenotypes of cells or non-human organisms, particularly of human cells or non-human mammals may be used in analytic procedures, e.g. in the functional and/or phenotypical analysis of complex physiological processes such as analysis of gene expression profiles and/or proteomes. For example, one may prepare the knock-out phenotypes of human genes in cultured cells which are assumed to be regulators of alternative splicing processes. Among these genes are particularly the members of the SR splicing factor family, e.g. ASF/SF2, SC35, SRp20, SRp40 or SRp55. Further, the effect of SR proteins on the mRNA profiles of predetermined alternatively spliced genes such as CD44 may be analysed. Preferably the analysis is carried out by high-throughput methods using oligonucleotide based chips.

Using RNAi based knockout technologies, the expression of an endogeneous target gene may be inhibited in a target cell or a target organism.

The endogeneous gene may be complemented by an exogeneous target nucleic acid coding for the target protein or a variant or mutated form of the target protein, e.g. a gene or a cDNA, which may optionally be fused to a further nucleic acid sequence encoding a detectable peptide or polypeptide, e.g. an affinity tag, particularly a multiple affinity tag. Variants or mutated forms of the target gene differ from the endogeneous target gene in that they encode a gene product which differs from the endogeneous gene product on the amino acid level by substitutions, insertions and/or deletions of single or multiple amino acids. The variants or mutated forms may have the same biological activity as the endogeneous target gene. On the other hand, the variant or mutated target gene may also have a biological activity, which differs from the biological activity of the endogeneous target gene, e.g. a partially deleted activity, a completely deleted activity, an enhanced activity etc.

The complementation may be accomplished by coexpressing the polypeptide encoded by the exogeneous nucleic acid, e.g. a fusion protein comprising the target protein and the affinity tag and the double stranded RNA molecule for knocking out the endogeneous gene in the target cell. This coexpression may be accomplished by using a suitable expression vector expressing both the polypeptide encoded by the exogeneous nucleic acid, e.g. the tag-modified target protein and the double stranded RNA molecule or alternatively by using a combination of expression vectors. Proteins and protein complexes which are synthesized de novo in the target cell will contain the exogeneous gene product, e.g. the modified fusion protein. In order to avoid suppression of the exogeneous gene product expression by the RNAi duplex molecule, the nucleotide sequence encoding the exogeneous nucleic acid may be altered on the DNA level (with or without causing mutations on the amino acid level) in the part of the sequence which is homologous to the double stranded RNA molecule. Alternatively, the endogeneous target gene may be complemented by corresponding nucleotide sequences from other species, e.g. from mouse.

Preferred applications for the cell or organism of the invention is the analysis of gene expression profiles and/or proteomes. In an especially preferred embodiment an analysis of a variant or mutant form of one or several target proteins is carried out, wherein said variant or mutant forms are reintroduced into the cell or organism by an exogeneous target nucleic acid as described above. The combination of knockout of an endogeneous gene and rescue by using mutated, e.g. partially deleted exogeneous target has advantages compared to the use of a knockout cell. Further, this method is particularly suitable for identifying functional domains of the target protein. In a further preferred embodiment a comparison, e.g. of gene expression profiles and/or proteomes and/or phenotypic characteristics of at least two cells or organisms is carried out. These organisms are selected from:

(i) a control cell or control organism without target gene inhibition,

(ii) a cell or organism with target gene inhibition and

(iii) a cell or organism with target gene inhibition plus target gene complementation by an exogeneous target nucleic acid.

The method and cell of the invention are also suitable in a procedure for identifying and/or characterizing pharmacological agents, e.g. identifying new pharmacological agents from a collection of test substances and/or characterizing mechanisms of action and/or side effects of known pharmacological agents.

Thus, the present invention also relates to a system for identifying and/or characterizing pharmacological agents acting on at least one target protein comprising:

(a) a eukaryotic cell or a eukaryotic non-human organism capable of expressing at least one endogeneous target gene coding for said target protein,

(b) at least one double-stranded RNA molecule capable of inhibiting the expression of said at least one endogeneous target gene, and

(c) a test substance or a collection of test substances wherein pharmacological properties of said test substance or said collection are to be identified and/or characterized.

Further, the system as described above preferably comprises:

(d) at least one exogeneous target nucleic acid coding for the target protein or a variant or mutated form of the target protein wherein said exogeneous target nucleic acid differs from the endogeneous target gene on the nucleic acid level such that the expression of the exogeneous target nucleic acid is substantially less inhibited by the double stranded RNA molecule than the expression of the endogeneous target gene.

Furthermore, the RNA knockout complementation method may be used for preparative purposes, e.g. for the affinity purification of proteins or protein complexes from eukaryotic cells, particularly mammalian cells and more particularly human cells. In this embodiment of the invention, the exogeneous target nucleic acid preferably codes for a target protein which is fused to an affinity tag.

The preparative method may be employed for the purification of high molecular weight protein complexes which preferably have a mass of ≥ 150 kD and more preferably of ≥ 500 kD and which optionally may contain nucleic acids such as RNA. Specific examples are the heterotrimeric protein complex consisting of the 20 kD, 60 kD and 90 kD proteins of the U4/U6 snRNP particle, the splicing factor SF3b from the 17S U2 snRNP consisting of 5 proteins having molecular weights of 14, 49, 120, 145 and 155 kD and the 25S U4/U6/U5 tri-snRNP particle containing the U4, U5 and U6 snRNA molecules and about 30 proteins, which has a molecular weight of about 1.7 MD.

This method is suitable for functional proteome analysis in mammalian cells, particularly human cells.

Further, the present invention is explained in more detail in the following figures and examples.

Figure Legends

Figure 1: Double-stranded RNA as short as 38 bp can mediate RNAi. (A) Graphic representation of dsRNAs used for targeting Pp-luc mRNA. Three series of blunt-ended dsRNAs covering a range of 29 to 504 bp were prepared. The position of the first nucleotide of the sense strand of the dsRNA is indicated relative to the start codon of Pp-luc mRNA (p1). (B) RNA interference assay (Tuschl et al., 1999). Ratios of target Pp-luc to control Rr-luc activity were normalized to a buffer control (black bar). DsRNAs (5 nM) were preincubated in Drosophila lysate for 15 min at 25°C prior to the addition of 7-methyl-guanosine-capped Pp-luc and Rr-luc mRNAs (~50 pM). The incubation was continued for another hour and then analyzed by the dual luciferase assay (Promega). The data are the average from at least four independent experiments ± standard deviation.

Figure 2: A 29 bp dsRNA is no longer processed to 21-23 nt fragments. Time course of 21-23 mer formation from processing of internally ^{32}P-labeled dsRNAs (5 nM) in the Drosophila lysate. The length and source of the dsRNA are indicated. An RNA size marker (M) has been loaded in the left lane and the fragment sizes are indicated. Double bands at time zero are due to incompletely denatured dsRNA.

Figure 3: Short dsRNAs cleave the mRNA target only once. (A) Denaturing gel electrophoreses of the stable 5' cleavage products produced by 1 h incubation of 10 nM sense or antisense RNA ^{32}P-labeled at the cap with 10 nM dsRNAs of the p133 series in Drosophila lysate.

Length markers were generated by partial nuclease T1 digestion and partial alkaline hydrolysis (OH) of the cap-labeled target RNA. The regions targeted by the dsRNAs are indicated as black bars on both sides. The 20-23 nt spacing between the predominant cleavage sites for the 111 bp long dsRNA is shown. The horizontal arrow indicates unspecific cleavage not due to RNAi. (B) Position of the cleavage sites on sense and antisense target RNAs. The sequences of the capped 177 nt sense and 180 nt antisense target RNAs are represented in antiparallel orientation such that complementary sequence are opposing each other. The region targeted by the different dsRNAs are indicated by differently colored bars positioned between sense and antisense target sequences. Cleavage sites are indicated by circles: large circle for strong cleavage, small circle for weak cleavage. The ^{32}P-radiolabeled phosphate group is marked by an asterisk.

Figure 4: 21 and 22 nt RNA fragments are generated by an RNase III-like mechanism.

(A) Sequences of ~21 nt RNAs after dsRNA processing. The ~21 nt RNA fragments generated by dsRNA processing were directionally cloned and sequenced. Oligoribonucleotides originating from the sense strand of the dsRNA are indicated as blue lines, those originating from the antisense strand as red lines. Thick bars are used if the same sequence was present in multiple clones, the number at the right indicating the frequency. The target RNA cleavage sites mediated by the dsRNA are indicated as orange circles, large circle for strong cleavage, small circle for weak cleavage (see Figure 3B). Circles on top of the sense strand indicated cleavage sites within the sense target and circles at the bottom of the dsRNA indicate cleavage site in the antisense target. Up to five additional nucleotides were identified in ~21 nt fragments derived from the 3' ends of the dsRNA. These nucleotides are random combinations of predominantly C, G, or A residues and were most likely added in an untemplated fashion during T7 transcription of the dsRNA-constituting strands. (B) Two-dimensional TLC analysis of the nucleotide composition of ~21 nt RNAs. The ~21 nt RNAs

were generated by incubation of internally radiolabeled 504 bp Pp-luc dsRNA in Drosophila lysate, gel-purified, and then digested to mononucleotides with nuclease P1 (top row) or ribonuclease T2 (bottom row). The dsRNA was internally radiolabeled by transcription in the presence of one of the indicated α-^{32}P nucleoside triphosphates. Radioactivity was detected by phosphorimaging. Nucleoside 5'-monophosphates, nucleoside 3'-monophosphates, nucleoside 5',3'-diphosphates, and inorganic phosphate are indicated as pN, Np, pNp, and p_i, respectively. Black circles indicate UV-absorbing spots from non-radioactive carrier nucleotides. The 3',5'-bisphosphates (red circles) were identified by co-migration with radiolabeled standards prepared by 5'-phosphorylation of nucleoside 3'-monophosphates with T4 polynucleotide kinase and γ-^{32}P-ATP.

Figure 5: Synthetic 21 and 22 nt RNAs Mediate Target RNA Cleavage. (A) Graphic representation of control 52 bp dsRNA and synthetic 21 and 22 nt dsRNAs. The sense strand of 21 and 22 nt short interfering RNAs (siRNAs) is shown blue, the antisense strand in red. The sequences of the siRNAs were derived from the cloned fragments of 52 and 111 bp dsRNAs (Figure 4A), except for the 22 nt antisense strand of duplex 5. The siRNAs in duplex 6 and 7 were unique to the 111 bp dsRNA processing reaction. The two 3' overhanging nucleotides indicated in green are present in the sequence of the synthetic antisense strand of duplexes 1 and 3. Both strands of the control 52 bp dsRNA were prepared by in vitro transcription and a fraction of transcripts may contain untemplated 3' nucleotide addition. The target RNA cleavage sites directed by the siRNA duplexes are indicated as orange circles (see legend to Figure 4A) and were determined as shown in Figure 5B. (B) Position of the cleavage sites on sense and antisense target RNAs. The target RNA sequences are as described in Figure 3B. Control 52 bp dsRNA (10 nM) or 21 and 22 nt RNA duplexes 1-7 (100 nM) were incubated with target RNA for 2.5 h at 25°C in Drosophila lysate. The stable 5' cleavage products were resolved on the gel. The cleavage sites are indicated in Figure 5A. The region targeted by the 52 bp

dsRNA or the sense (s) or antisense (as) strands are indicated by the black bars to the side of the gel. The cleavage sites are all located within the region of identity of the dsRNAs. For precise determination of the cleavage sites of the antisense strand, a lower percentage gel was used.

Figure 6: Long 3' overhangs on short dsRNAs inhibit RNAi.
(A) Graphic representation of 52 bp dsRNA constructs. The 3' extensions of sense and antisense strand are indicated in blue and red, respectively. The observed cleavage sites on the target RNAs are represented as orange circles analogous to Figure 4A and were determined as shown in Figure 6B. (B) Position of the cleavage sites on sense and antisense target RNAs. The target RNA sequences are as described in Figure 3B. DsRNA (10 nM) was incubated with target RNA for 2.5 h at 25°C in Drosophila lysate. The stable 5' cleavage products were resolved on the gel. The major cleavage sites are indicated with a horizontal arrow and also represented in Figure 6A. The region targeted by the 52 bp dsRNA is represented as black bar at both sides of the gel.

Figure 7: Proposed Model for RNAi.
RNAi is predicted to begin with processing of dsRNA (sense strand in black, antisense strand in red) to predominantly 21 and 22 nt short interfering RNAs (siRNAs). Short overhanging 3' nucleotides, if present on the dsRNA, may be beneficial for processing of short dsRNAs. The dsRNA-processing proteins, which remain to be characterized, are represented as green and blue ovals, and assembled on the dsRNA in asymmetric fashion. In our model, this is illustrated by binding of a hypothetical blue protein or protein domain with the siRNA strand in 3' to 5' direction while the hypothetical green protein or protein domain is always bound to the opposing siRNA strand. These proteins or a subset remain associated with the siRNA duplex and preserve its orientation as determined by the direction of the dsRNA processing reaction. Only the siRNA sequence associated with the blue protein is able to guide target RNA cleavage. The endonuclease com-

plex is referred to as small interfering ribonucleoprotein complex or siRNP. It is presumed here, that the endonuclease that cleaves the dsRNA may also cleave the target RNA, probably by temporarily displacing the passive siRNA strand not used for target recognition. The target RNA is then cleaved in the center of the region recognized by the sequence-complementary guide siRNA.

Figure 8: Reporter constructs and siRNA duplexes.
(a) The firefly (Pp-luc) and sea pansy (Rr-luc) luciferase reporter gene regions from plasmids pGL2-Control, pGL-3-Control and pRL-TK (Promega) are illustrated. SV40 regulatory elements, the HSV thymidine kinase promoter and two introns (lines) are indicated. The sequence of GL3 luciferase is 95% identical to GL2, but RL is completely unrelated to both. Luciferase expression from pGL2 is approx. 10-fold lower than from pGL3 in transfected mammalian cells. The region targeted by the siRNA duplexes is indicated as black bar below the coding region of the luciferase genes. (b) The sense (top) and antisense (bottom) sequences of the siRNA duplexes targeting GL2, GL3 and RL luciferase are shown. The GL2 and GL3 siRNA duplexes differ by only 3 single nucleotide substitutions (boxed in gray). As unspecific control, a duplex with the inverted GL2 sequence, invGL2, was synthesized. The 2 nt 3' overhang of 2'-deoxythymidine is indicated as TT; uGL2 is similar to GL2 siRNA but contains ribo-uridine 3' overhangs.

Figure 9: RNA interference by siRNA duplexes.
Ratios of target control luciferase were normalized to a buffer control (bu, black bars); gray bars indicate ratios of *Photinus pyralis* (Pp-luc) GL2 or GL3 luciferase to *Renilla reniformis* (Rr-luc) RL luciferase (left axis), white bars indicate RL to GL2 or GL3 ratios (right axis). Panels a, c, e, g and i describe experiments performed with the combination of pGL2-Control and pRL-TK reporter plasmids, panels b, d, f, h and j with pGL3-Control and pRL-TK reporter plasmids. The cell line used for the interference experiment is indicated at the top of each plot. The ratios of Pp-luc/Rr-luc for the

buffer control (bu) varied between 0.5 and 10 for pGL2/pRL and between 0.03 and 1 for pGL3/pRL, respectively, before normalization and between the various cell lines tested. The plotted data were averaged from three independent experiments ± S.D.

Figure 10: Effects of 21 nt siRNA, 50 bp and 500 bp dsRNAs on luciferase expression in HeLa cells.

The exact length of the long dsRNAs is indicated below the bars. Panels a, c and e describe experiments performed with pGL2-Control and pRL-TK reporter plasmids, panels b, d and f with pGL3-Control and pRL-TK reporter plasmids. The data were averaged from two independent experiments ± S.D. (a), (b) Absolute Pp-luc expression, plotted in arbitrary luminescence units. (c), (d) Rr-luc expression, plotted in arbitrary luminescence units. (e), (f) Ratios of normalized target to control luciferase. The ratios of luciferase activity for siRNA duplexes were normalized to a buffer control (bu, black bars); the luminescence ratios for 50 or 500 bp dsRNAs were normalized to the respective ratios observed for 50 and 500 bp dsRNA from humanized GFP (hG, black bars). It should be noted that the overall differences in sequences between the 49 and 484 bp dsRNAs targeting GL2 and GL3 are not sufficient to confer specificity between GL2 and GL3 targets (43 nt uninterrupted identity in 49 bp segment, 239 nt longest uninterrupted identity in 484 bp segment).

Figure 11: Variation of the 3' overhang of duplexes of 21-nt siRNAs.

(A) Outline of the experimental strategy. The capped and polyadenylated sense target mRNA is depicted and the relative positions of sense and antisense siRNAs are shown. Eight series of duplexes, according to the eight different antisense strands were prepared. The siRNA sequences and the number of overhanging nucleotides were changed in 1-nt steps. (B) Normalized relative luminescence of target luciferase (Photinus pyralis, Pp-luc) to control luciferase (Renilla reniformis, Rr-luc) in D. melanogaster embryo lysate in the presence of 5 nM blunt-ended dsRNAs. The lumi-

nescence ratios determined in the presence of dsRNA were normalized to the ratio obtained for a buffer control (bu, black bar). Normalized ratios less than 1 indicate specific interference. (C-J) Normalized interference ratios for eight series of 21-nt siRNA duplexes. The sequences of siRNA duplexes are depicted above the bar graphs. Each panel shows the interference ratio for a set of duplexes formed with a given antisense guide siRNA and 5 different sense siRNAs. The number of overhanging nucleotides (3' overhang, positive numbers; 5' overhangs, negative numbers) is indicated on the x-axis. Data points were averaged from at least 3 independent experiments, error bars represent standard deviations.

Figure 12: Variation of the length of the sense strand of siRNA duplexes. (A) Graphic representation of the experiment. Three 21-nt antisense strands were paired with eight sense siRNAs. The siRNAs were changed in length at their 3' end. The 3' overhang of the antisense siRNA was 1-nt (B), 2-nt (C), or 3-nt (D) while the sense siRNA overhang was varied for each series. The sequences of the siRNA duplexes and the corresponding interference ratios are indicated.

Figure 13: Variation of the length of siRNA duplexes with preserved 2-nt 3' overhangs.
(A) Graphic representation of the experiment. The 21-nt siRNA duplex is identical in sequence to the one shown in Figure 11H or 12C. The siRNA duplexes were extended to the 3' side of the sense siRNA (B) or the 5' side of the sense siRNA (C). The siRNA duplex sequences and the respective interference ratios are indicated.

Figure 14: Substitution of the 2'-hydroxyl groups of the siRNA ribose residues.
The 2'-hydroxyl groups (OH) in the strands of siRNA duplexes were replaced by 2'-deoxy (d) or 2'-O-methyl (Me). 2-nt and 4-nt 2'-deoxy substitu-

tions at the 3'-ends are indicated as 2-nt d and 4-nt d, respectively. Uridine residues were replaced by 2'-deoxy thymidine.

Figure 15: Mapping of sense and antisense target RNA cleavage by 21-nt siRNA duplexes with 2-nt 3' overhangs.

(A) Graphic representation of 32P (asterisk) cap-labelled sense and antisense target RNAs and siRNA duplexes. The position of sense and antisense target RNA cleavage is indicated by triangles on top and below the siRNA duplexes, respectively. (B) Mapping of target RNA cleavage sites. After 2 h incubation of 10 nM target with 100 nM siRNA duplex in D. melanogaster embryo lysate, the 5' cap-labelled substrate and the 5' cleavage products were resolved on sequencing gels. Length markers were generated by partial RNase T1 digestion (T1) and partial alkaline hydrolysis (OH-) of the target RNAs. The bold lines to the left of the images indicate the region covered by the siRNA strands 1 and 5 of the same orientation as the target.

Figure 16: The 5' end of a guide siRNA defines the position of target RNA cleavage.

(A, B) Graphic representation of the experimental strategy. The antisense siRNA was the same in all siRNA duplexes, but the sense strand was varied between 18 to 25 nt by changing the 3' end (A) or 18 to 23 nt by changing the 5' end (B). The position of sense and antisense target RNA cleavage is indicated by triangles on top and below the siRNA duplexes, respectively. (C, D) Analysis of target RNA cleavage using cap-labelled sense (top panel) or antisense (bottom panel) target RNAs. Only the cap-labelled 5' cleavage products are shown. The sequences of the siRNA duplexes are indicated, and the length of the sense siRNA strands is marked on top of the panel. The control lane marked with a dash in panel (C) shows target RNA incubated in absence of siRNAs. Markers were as described in Figure 15. The arrows in (D), bottom panel, indicate the target RNA cleavage sites that differ by 1 nt.

Figure 17: Sequence variation of the 3' overhang of siRNA duplexes.
The 2-nt 3' overhang (NN, in gray) was changed in sequence and composition as indicated (T, 2'-deoxythymidine, dG, 2'-deoxyguanosine; asterisk, wild-type siRNA duplex). Normalized interference ratios were determined as described in Figure 11. The wild-type sequence is the same as depicted in Figure 14.

Figure 18: Sequence specificity of target recognition.
The sequences of the mismatched siRNA duplexes are shown, modified sequence segments or single nucleotides are underlayed in gray. The reference duplex (ref) and the siRNA duplexes 1 to 7 contain 2'-deoxythymidine 2-nt overhangs. The silencing efficiency of the thymidine-modified reference duplex was comparable to the wild-type sequence (Figure 17). Normalized interference ratios were determined as described in Figure 11.

Figure 19: Variation of the length of siRNA duplexes with preserved 2-nt 3' overhangs.
The siRNA duplexes were extended to the 3' side of the sense siRNA (A) or the 5' side of the sense siRNA (B). The siRNA duplex sequences and the respective interference ratios are indicated. For HeLa SS6 cells, siRNA duplexes (0.84 μg) targeting GL2 luciferase were transfected together with pGL2-Control and pRL-TK plasmids. For comparison, the in vitro RNAi activities of siRNA duplexes tested in D. melanogaster lysate are indicated.

Example 1

RNA Interference Mediated by Small Synthetic RNAs

1.1. Experimental Procedures

1.1.1 In Vitro RNAi

In vitro RNAi and lysate preparations were performed as described previously (Tuschl et al., 1999; Zamore et al., 2000). It is critical to use

freshly dissolved creatine kinase (Roche) for optimal ATP regeneration. The RNAi translation assays (Fig. 1) were performed with dsRNA concentrations of 5 nM and an extended pre-incubation period of 15 min at 25°C prior to the addition of in vitro transcribed, capped and polyadenylated Pp-luc and Rr-luc reporter mRNAs. The incubation was continued for 1 h and the relative amount of Pp-luc and Rr-luc protein was analyzed using the dual luciferase assay (Promega) and a Monolight 3010C luminometer (PharMingen).

1.1.2 RNA Synthesis

Standard procedures were used for in vitro transcription of RNA from PCR templates carrying T7 or SP6 promoter sequences, see for example (Tuschl et al., 1998). Synthetic RNA was prepared using Expedite RNA phosphoramidites (Proligo). The 3' adapter oligonucleotide was synthesized using dimethoxytrityl-1,4-benzenedimethanol-succinyl-aminopropyl-CPG. The oligoribonucleotides were deprotected in 3 ml of 32% ammonia/ethanol (3/1) for 4 h at 55°C (Expedite RNA) or 16 h at 55°C (3' and 5' adapter DNA/RNA chimeric oligonucleotides) and then desilylated and gel-purified as described previously (Tuschl et al., 1993). RNA transcripts for dsRNA preparation including long 3' overhangs were generated from PCR templates that contained a T7 promoter in sense and an SP6 promoter in antisense direction. The transcription template for sense and antisense target RNA was PCR-amplified with GCGTAATACGACTCACTATAGAACAATTGCTTTTACAG (underlined, T7 promoter) as 5' primer and ATTTAGGTGACACTATAGGCATAAAGAATTGAAGA (underlined, SP6 promoter) as 3' primer and the linearized Pp-luc plasmid (pGEM-luc sequence) (Tuschl et al., 1999) as template; the T7-transcribed sense RNA was 177 nt long with the Pp-luc sequence between pos. 113-273 relative to the start codon and followed by 17 nt of the complement of the SP6 promoter sequence at the 3' end. Transcripts for blunt-ended dsRNA

formation were prepared by transcription from two different PCR products which only contained a single promoter sequence.

DsRNA annealing was carried out using a phenol/chloroform extraction. Equimolar concentration of sense and antisense RNA (50 nM to 10 μM, depending on the length and amount available) in 0.3 M NaOAc (pH 6) were incubated for 30 s at 90°C and then extracted at room temperature with an equal volume of phenol/chloroform, and followed by a chloroform extraction to remove residual phenol. The resulting dsRNA was precipitated by addition of 2.5-3 volumes of ethanol. The pellet was dissolved in lysis buffer (100 mM KCl, 30 mM HEPES-KOH, pH 7.4, 2 mM $Mg(OAc)_2$) and the quality of the dsRNA was verified by standard agarose gel electrophoreses in 1 x TAE-buffer. The 52 bp dsRNAs with the 17 nt and 20 nt 3' overhangs (Figure 6) were annealed by incubating for 1 min at 95 °C, then rapidly cooled to 70°C and followed by slow cooling to room temperature over a 3 h period (50 μl annealing reaction, 1 μM strand concentration, 300 mM NaCl, 10 mM Tris-HCl, pH 7.5). The dsRNAs were then phenol/chloroform extracted, ethanol-precipitated and dissolved in lysis buffer.

Transcription of internally ^{32}P-radiolabeled RNA used for dsRNA preparation (Figures 2 and 4) was performed using 1 mM ATP, CTP, GTP, 0.1 or 0.2 mM UTP, and 0.2-0.3 μM -^{32}P-UTP (3000 Ci/mmol), or the respective ratio for radiolabeled nucleoside triphosphates other than UTP. Labeling of the cap of the target RNAs was performed as described previously. The target RNAs were gel-purified after cap-labeling.

1.1.3 Cleavage Site Mapping

Standard RNAi reactions were performed by pre-incubating 10 nM dsRNA for 15 min followed by addition of 10 nM cap-labeled target RNA. The reaction was stopped after a further 2 h (Figure 2A) or 2.5 h incubation (Figure 5B and 6B) by proteinase K treatment (Tuschl et al., 1999). The samples were then analyzed on 8 or 10% sequencing gels. The 21 and 22

nt synthetic RNA duplexes were used at 100 nM final concentration (Fig 5B).

1.1.4 Cloning of ~21 nt RNAs

The 21 nt RNAs were produced by incubation of radiolabeled dsRNA in Drosophila lysate in absence of target RNA (200 µl reaction, 1 h incubation, 50 nM dsP111, or 100 nM dsP52 or dsP39). The reaction mixture was subsequently treated with proteinase K (Tuschl et al., 1999) and the dsRNA-processing products were separated on a denaturing 15% polyacrylamide gel. A band, including a size range of at least 18 to 24 nt, was excised, eluted into 0.3 M NaCl overnight at 4°C and in siliconized tubes. The RNA was recovered by ethanol-precipitation and dephosphorylated (30 µl reaction, 30 min, 50°C, 10 U alkaline phosphatase, Roche). The reaction was stopped by phenol/chloroform extraction and the RNA was ethanol-precipitated. The 3' adapter oligonucleotide (pUUUaaccgcatccttctcx: uppercase, RNA; lowercase, DNA; p, phosphate; x, 4-hydroxymethylbenzyl) was then ligated to the dephosphorylated ~21 nt RNA (20 µl reaction, 30 min, 37°C, 5 µM 3' adapter, 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 0.2 mM ATP, 0.1 mg/ml acetylated BSA, 15% DMSO, 25 U T4 RNA ligase, Amersham-Pharmacia) (Pan and Uhlenbeck, 1992). The ligation reaction was stopped by the addition of an equal volume of 8 M urea/50 mM EDTA stopmix and directly loaded on a 15% gel. Ligation yields were greater 50%. The ligation product was recovered from the gel and 5'-phosphorylated (20 µl reaction, 30 min, 37°C, 2 mM ATP, 5 U T4 polynucleotide kinase, NEB). The phosphorylation reaction was stopped by phenol/chloroform extraction and RNA was recovered by ethanol-precipitation. Next, the 5' adapter (tactaatacgactcactAAA: uppercase, RNA; lowercase, DNA) was ligated to the phosphorylated ligation product as described above. The new ligation product was gel-purified and eluted from the gel slice in the presence of reverse transcription primer (GACTAGCTGGAATTCAAGGATGCGGTTAAA: bold, Eco RI site) used as carrier. Reverse transcription (15 µl reaction, 30 min, 42°C, 150 U Super-

script II reverse transcriptase, Life Technologies) was followed by PCR using as 5' primer CAGCCAACG**GAATTC**ATACGACTCACTAAA (bold, Eco RI site) and the 3' RT primer. The PCR product was purified by phenol/ chloroform extraction and ethanol-precipitated. The PCR product was then digested with Eco RI (NEB) and concatamerized using T4 DNA ligase (high conc., NEB). Concatamers of a size range of 200 to 800 bp were separated on a low-melt agarose gel, recovered from the gel by a standard melting and phenol extraction procedure, and ethanol-precipitated. The unpaired ends were filled in by incubation with Taq polymerase under standard conditions for 15 min at 72°C and the DNA product was directly ligated into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen). Colonies were screened using PCR and M13-20 and M13 Reverse sequencing primers. PCR products were directly submitted for custom sequencing (Sequence Laboratories Göttingen GmbH, Germany). On average, four to five 21mer sequences were obtained per clone.

1.1.5 2D-TLC Analysis

Nuclease P1 digestion of radiolabeled, gel-purified siRNAs and 2D-TLC was carried out as described (Zamore et al., 2000). Nuclease T2 digestion was performed in 10 µl reactions for 3 h at 50°C in 10 mM ammonium acetate (pH 4.5) using 2 µg/µl carrier tRNA and 30 U ribonuclease T2 (Life Technologies). The migration of non-radioactive standards was determined by UV shadowing. The identity of nucleoside-3',5'-disphosphates was confirmed by co-migration of the T2 digestion products with standards prepared by 5'-^{32}P-phosphorylation of commercial nucleoside 3'-monophosphates using γ-32P-ATP and T4 polynucleotide kinase (data not shown).

1.2 Results and Discussion

1.2.1 Length Requirements for Processing of dsRNA to 21 and 22 nt RNA Fragments

Lysate prepared from D. melanogaster syncytial embryos recapitulates RNAi in vitro providing a novel tool for biochemical analysis of the mechanism of RNAi (Tuschl et al., 1999; Zamore et al., 2000). In vitro and in vivo analysis of the length requirements of dsRNA for RNAi has revealed that short dsRNA (<150 bp) are less effective than longer dsRNAs in degrading target mRNA (Caplen et al., 2000; Hammond et al., 2000; Ngo et al., 1998); Tuschl et al., 1999). The reasons for reduction in mRNA degrading efficiency are not understood. We therefore examined the precise length requirement of dsRNA for target RNA degradation under optimized conditions in the Drosophila lysate (Zamore et al., 2000). Several series of dsRNAs were synthesized and directed against firefly luciferase (Pp-luc) reporter RNA. The specific suppression of target RNA expression was monitored by the dual luciferase assay (Tuschl et al., 1999) (Figures 1A and 1B). We detected specific inhibition of target RNA expression for dsRNAs as short as 38 bp, but dsRNAs of 29 to 36 bp were not effective in this process. The effect was independent of the target position and the degree of inhibition of Pp-luc mRNA expression correlated with the length of the dsRNA, i.e. long dsRNAs were more effective than short dsRNAs.

It has been suggested that the 21-23 nt RNA fragments generated by processing of dsRNAs are the mediators of RNA interference and co-suppression (Hamilton and Baulcombe, 1999; Hammond et al., 2000; Zamore et al., 2000). We therefore analyzed the rate of 21-23 nt fragment formation for a subset of dsRNAs ranging in size between 501 to 29 bp. Formation of 21-23 nt fragments in Drosophila lysate (Figure 2) was readily detectable for 39 to 501 bp long dsRNAs but was significantly delayed for the 29 bp dsRNA. This observation is consistent with a role of 21-23 nt fragments in guiding mRNA cleavage and provides an explanation for the

lack of RNAi by 30 bp dsRNAs. The length dependence of 21-23 mer formation is likely to reflect a biologically relevant control mechanism to prevent the undesired activation of RNAi by short intramolecular base-paired structures of regular cellular RNAs.

1.2.2 39 bp dsRNA Mediates Target RNA Cleavage at a Single Site

Addition of dsRNA and 5'-capped target RNA to the Drosophila lysate results in sequence-specific degradation of the target RNA (Tuschl et al., 1999). The target mRNA is only cleaved within the region of identity with the dsRNA and many of the target cleavage sites were separated by 21-23 nt (Zamore et al., 2000). Thus, the number of cleavage sites for a given dsRNA was expected to roughly correspond to the length of the dsRNA divided by 21. We mapped the target cleavage sites on a sense and an antisense target RNA which was 5' radiolabeled at the cap (Zamore et al., 2000) (Figures 3A and 3B). Stable 5' cleavage products were separated on a sequencing gel and the position of cleavage was determined by comparison with a partial RNase T1 and an alkaline hydrolysis ladder from the target RNA.

Consistent with the previous observation (Zamore et al., 2000), all target RNA cleavage sites were located within the region of identity to the dsRNA. The sense or the antisense traget was only cleaved once by 39 bp dsRNA. Each cleavage site was located 10 nt from the 5' end of the region covered by the dsRNA (Figure 3B). The 52 bp dsRNA, which shares the same 5' end with the 39 bp dsRNA, produces the same cleavage site on the sense target, located 10 nt from the 5' end of the region of identity with the dsRNA, in addition to two weaker cleavage sites 23 and 24 nt downstream of the first site. The antisense target was only cleaved once, again 10 nt from the 5' end of the region covered by its respective dsRNA. Mapping of the cleavage sites for the 38 to 49 bp dsRNAs shown in Figure 1 showed that the first and predominant cleavage site was always located 7 to 10 nt downstream of the region covered by the dsRNA (data not

shown). This suggests that the point of target RNA cleavage is determined by the end of the dsRNA and could imply that processing to 21-23 mers starts from the ends of the duplex.

Cleavage sites on sense and antisense target for the longer 111 bp dsRNA were much more frequent than anticipated and most of them appear in clusters separated by 20 to 23 nt (Figures 3A and 3B). As for the shorter dsRNAs, the first cleavage site on the sense target is 10 nt from the 5′ end of the region spanned by the dsRNA; and the first cleavage site on the antisense target is located 9 nt from the 5′ end of region covered by the dsRNA. It is unclear what causes this disordered cleavage, but one possibility could be that longer dsRNAs may not only get processed from the ends but also internally, or there are some specificity determinants for dsRNA processing which we do not yet understand. Some irregularities to the 21-23 nt spacing were also previously noted (Zamore et al., 2000). To better understand the molecular basis of dsRNA processing and target RNA recognition, we decided to analyze the sequences of the 21-23 nt fragments generated by processing of 39, 52, and 111 bp dsRNAs in the Drosophila lysate.

1.2.3 dsRNA is Processed to 21 and 22 nt RNAs by an RNase III-Like Mechanism

In order to characterize the 21-23 nt RNA fragments we examined the 5′ and 3′ termini of the RNA fragments. Periodate oxidation of gel-purified 21-23 nt RNAs followed by ß-elimination indicated the presence of a terminal 2′ and 3′ hydroxyl groups. The 21-23 mers were also responsive to alkaline phosphatase treatment indicating the presence of a 5′ terminal phosphate group. The presence of 5′ phosphate and 3′ hydroxyl termini suggests that the dsRNA could be processed by an enzymatic activity similar to E. coli RNase III (for reviews, see (Dunn, 1982; Nicholson, 1999; Robertson, 1990; Robertson, 1982)).

Directional cloning of 21-23 nt RNA fragments was performed by ligation of a 3' and 5' adapter oligonucleotide to the purified 21-23 mers using T4 RNA ligase. The ligation products were reverse transcribed, PCR-amplified, concatamerized, cloned, and sequenced. Over 220 short RNAs were sequenced from dsRNA processing reactions of the 39, 52 and 111 bp dsRNAs (Figure 4A). We found the following length distribution: 1% 18 nt, 5% 19 nt, 12% 20 nt, 45% 21 nt, 28% 22 nt, 6% 23 nt, and 2% 24 nt. Sequence analysis of the 5' terminal nucleotide of the processed fragments indicated that oligonucleotides with a 5' guanosine were underrepresented. This bias was most likely introduced by T4 RNA ligase which discriminates against 5' phosphorylated guanosine as donor oligonucleotide; no significant sequence bias was seen at the 3' end. Many of the ~21 nt fragments derived from the 3' ends of the sense or antisense strand of the duplexes include 3' nucleotides that are derived from untemplated addition of nucleotides during RNA synthesis using T7 RNA polymerase. Interestingly, a significant number of endogenous Drosophila ~21 nt RNAs were also cloned, some of them from LTR and non-LTR retrotransposons (data not shown). This is consistent with a possible role for RNAi in transposon silencing.

The ~21 nt RNAs appear in clustered groups (Figure 4A) which cover the entire dsRNA sequences. Apparently, the processing reaction cuts the dsRNA by leaving staggered 3' ends, another characteristic of RNase III cleavage. For the 39 bp dsRNA, two clusters of ~21 nt RNAs were found from each dsRNA-constituting strand including overhanging 3' ends, yet only one cleavage site was detected on the sense and antisense target (Figures 3A and 3B). If the ~21 nt fragments were present as single-stranded guide RNAs in a complex that mediates mRNA degradation, it could be assumed that at least two target cleavage sites exist, but this was not the case. This suggests that the ~21 nt RNAs may be present in double-stranded form in the endonuclease complex but that only one of the strands can be used for target RNA recognition and cleavage. The use of

only one of the ~21 nt strands for target cleavage may simply be determined by the orientation in which the ~21 nt duplex is bound to the nuclease complex. This orientation is defined by the direction in which the original dsRNA was processed.

The ~21mer clusters for the 52 bp and 111 bp dsRNA are less well defined when compared to the 39 bp dsRNA. The clusters are spread over regions of 25 to 30 nt most likely representing several distinct subpopulations of ~21 nt duplexes and therefore guiding target cleavage at several nearby sites. These cleavage regions are still predominantly separated by 20 to 23 nt intervals. The rules determining how regular dsRNA can be processed to ~21 nt fragments are not yet understood, but it was previously observed that the approx. 21-23 nt spacing of cleavage sites could be altered by a run of uridines (Zamore et al., 2000). The specificity of dsRNA cleavage by E. coli RNase III appears to be mainly controlled by antideterminants, i.e. excluding some specific base-pairs at given positions relative to the cleavage site (Zhang and Nicholson, 1997).

To test whether sugar-, base- or cap-modification were present in processed ~21 nt RNA fragments, we incubated radiolabeled 505 bp Pp-luc dsRNA in lysate for 1 h, isolated the ~21 nt products, and digested it with P1 or T2 nuclease to mononucleotides. The nucleotide mixture was then analyzed by 2D thin-layer chromatography (Figure 4B). None of the four natural ribonucleotides were modified as indicated by P1 or T2 digestion. We have previously analyzed adenosine to inosine conversion in the ~21 nt fragments (after a 2 h incubation) and detected a small extent (<0.7%) deamination (Zamore et al., 2000); shorter incubation in lysate (1 h) reduced this inosine fraction to barely detectable levels. RNase T2, which cleaves 3' of the phosphodiester linkage, produced nucleoside 3'-phosphate and nucleoside 3',5'-diphosphate, thereby indicating the presence of a 5'-terminal monophosphate. All four nucleoside 3',5'-diphosphates were detected and suggest that the internucleotidic linkage was

cleaved with little or no sequence-specificity. In summary, the ~21 nt fragments are unmodified and were generated from dsRNA such that 5'-monophosphates and 3'-hydroxyls were present at the 5'-end.

1.2.4 Synthetic 21 and 22 nt RNAs Mediate Target RNA Cleavage

Analysis of the products of dsRNA processing indicated that the ~21 nt fragments are generated by a reaction with all the characteristics of an RNase III cleavage reaction (Dunn, 1982; Nicholson, 1999; Robertson, 1990; Robertson, 1982). RNase III makes two staggered cuts in both strands of the dsRNA, leaving a 3' overhang of about 2 nt. We chemically synthesized 21 and 22 nt RNAs, identical in sequence to some of the cloned ~21 nt fragments, and tested them for their ability to mediate target RNA degradation (Figures 5A and 5B). The 21 and 22 nt RNA duplexes were incubated at 100 nM concentrations in the lysate, a 10-fold higher concentrations than the 52 bp control dsRNA. Under these conditions, target RNA cleavage is readily detectable. Reducing the concentration of 21 and 22 nt duplexes from 100 to 10 nM does still cause target RNA cleavage. Increasing the duplex concentration from 100 nM to 1000 nM however does not further increase target cleavage, probably due to a limiting protein factor within the lysate.

In contrast to 29 or 30 bp dsRNAs that did not mediate RNAi, the 21 and 22 nt dsRNAs with overhanging 3' ends of 2 to 4 nt mediated efficient degradation of target RNA (duplexes 1, 3, 4, 6, Figures 5A and 5B). Blunt-ended 21 or 22 nt dsRNAs (duplexes 2, 5, and 7, Figures 5A and 5B) were reduced in their ability to degrade the target and indicate that overhanging 3' ends are critical for reconstitution of the RNA-protein nuclease complex. The single-stranded overhangs may be required for high affinity binding of the ~ 21 nt duplex to the protein components. A 5' terminal phosphate, although present after dsRNA processing, was not required to mediate target RNA cleavage and was absent from the short synthetic RNAs.

The synthetic 21 and 22 nt duplexes guided cleavage of sense as well as antisense targets within the region covered by the short duplex. This is an important result considering that a 39 bp dsRNA, which forms two pairs of clusters of ~21 nt fragments (Fig. 2), cleaved sense or antisense target only once and not twice. We interpret this result by suggesting that only one of two strands present in the ~21 nt duplex is able to guide target RNA cleavage and that the orientation of the ~21 nt duplex in the nuclease complex is determined by the initial direction of dsRNA processing. The presentation of an already perfectly processed ~21 nt duplex to the in vitro system however does allow formation of the active sequence-specific nuclease complex with two possible orientations of the symmetric RNA duplex. This results in cleavage of sense as well as antisense target within the region of identity with the 21 nt RNA duplex.

The target cleavage site is located 11 or 12 nt downstream of the first nucleotide that is complementary to the 21 or 22 nt guide sequence, i.e. the cleavage site is near center of the region covered by the 21 or 22 nt RNAs (Figures 4A and 4B). Displacing the sense strand of a 22 nt duplex by two nucleotides (compare duplexes 1 and 3 in Figure 5A) displaced the cleavage site of only the antisense target by two nucleotides. Displacing both sense and antisense strand by two nucleotides shifted both cleavage sites by two nucleotides (compare duplexes 1 and 4). We predict that it will be possible to design a pair of 21 or 22 nt RNAs to cleave a target RNA at almost any given position.

The specificity of target RNA cleavage guided by 21 and 22 nt RNAs appears exquisite as no aberrant cleavage sites are detected (Figure 5B). It should however be noted, that the nucleotides present in the 3' overhang of the 21 and 22 nt RNA duplex may contribute less to substrate recognition than the nucleotides near the cleavage site. This is based on the observation that the 3' most nucleotide in the 3' overhang of the active duplexes 1 or 3 (Figure 5A) is not complementary to the target. A detailed

analysis of the specificity of RNAi can now be readily undertaken using synthetic 21 and 22 nt RNAs.

Based on the evidence that synthetic 21 and 22 nt RNAs with overhanging 3' ends mediate RNA interference, we propose to name the ~21 nt RNAs "short interfering RNAs" or siRNAs and the respective RNA-protein complex a "small interfering ribonucleoprotein particle" or siRNP.

1.2.5 3' Overhangs of 20 nt on short dsRNAs inhibit RNAi

We have shown that short blunt-ended dsRNAs appear to be processed from the ends of the dsRNA. During our study of the length dependence of dsRNA in RNAi, we have also analyzed dsRNAs with 17 to 20 nt overhanging 3' ends and found to our surprise that they were less potent than blunt-ended dsRNAs. The inhibitory effect of long 3' ends was particularly pronounced for dsRNAs up to 100 bp but was less dramatic for longer dsRNAs. The effect was not due to imperfect dsRNA formation based on native gel analysis (data not shown). We tested if the inhibitory effect of long overhanging 3' ends could be used as a tool to direct dsRNA processing to only one of the two ends of a short RNA duplex.

We synthesized four combinations of the 52 bp model dsRNA, blunt-ended, 3' extension on only the sense strand, 3'-extension on only the antisense strand, and double 3' extension on both strands, and mapped the target RNA cleavage sites after incubation in lysate (Figures 6A and 6B). The first and predominant cleavage site of the sense target was lost when the 3' end of the antisense strand of the duplex was extended, and vice versa, the strong cleavage site of the antisense target was lost when the 3' end of sense strand of the duplex was extended. 3' Extensions on both strands rendered the 52 bp dsRNA virtually inactive. One explanation for the dsRNA inactivation by ~20 nt 3' extensions could be the association of single-stranded RNA-binding proteins which could interfere with the association of one of the dsRNA-processing factors at this end. This result

is also consistent with our model where only one of the strands of the siRNA duplex in the assembled siRNP is able to guide target RNA cleavage. The orientation of the strand that guides RNA cleavage is defined by the direction of the dsRNA processing reaction. It is likely that the presence of 3′ staggered ends may facilitate the assembly of the processing complex. A block at the 3′ end of the sense strand will only permit dsRNA processing from the opposing 3′ end of the antisense strand. This in turn generates siRNP complexes in which only the antisense strand of the siRNA duplex is able to guide sense target RNA cleavage. The same is true for the reciprocal situation.

The less pronounced inhibitory effect of long 3′ extensions in the case of longer dsRNAs (≥500 bp, data not shown) suggests to us that long dsRNAs may also contain internal dsRNA-processing signals or may get processed cooperatively due to the association of multiple cleavage factors.

1.2.6 A Model for dsRNA-Directed mRNA Cleavage

The new biochemical data update the model for how dsRNA targets mRNA for destruction (Figure 7). Double-stranded RNA is first processed to short RNA duplexes of predominantly 21 and 22 nt in length and with staggered 3′ ends similar to an RNase III-like reaction (Dunn, 1982; Nicholson, 1999; Robertson, 1982). Based on the 21-23 nt length of the processed RNA fragments it has already been speculated that an RNase III-like activity may be involved in RNAi (Bass, 2000). This hypothesis is further supported by the presence of 5′ phosphates and 3′ hydroxyls at the termini of the siRNAs as observed in RNase III reaction products (Dunn, 1982; Nicholson, 1999). Bacterial RNase III and the eukaryotic homologs Rnt1p in S. cerevisiae and Pac1p in S. pombe have been shown to function in processing of ribosomal RNA as well as snRNA and snoRNAs (see for example Chanfreau et al., 2000).

Little is known about the biochemistry of RNase III homologs from plants, animals or human. Two families of RNase III enzymes have been identified predominantly by database-guided sequence analysis or cloning of cDNAs. The first RNase III family is represented by the 1327 amino acid long D. melanogaster protein drosha (Acc. AF116572). The C-terminus is composed of two RNase III and one dsRNA-binding domain and the N-terminus is of unknown function. Close homologs are also found in C. elegans (Acc. AF160248) and human (Acc. AF189011) (Filippov et al., 2000; Wu et al., 2000). The drosha-like human RNase III was recently cloned and characterized (Wu et al., 2000). The gene is ubiquitously expressed in human tissues and cell lines, and the protein is localized in the nucleus and the nucleolus of the cell. Based on results inferred from antisense inhibition studies, a role of this protein for rRNA processing was suggested. The second class is represented by the C. elegans gene K12H4.8 (Acc. S44849) coding for a 1822 amino acid long protein. This protein has an N-terminal RNA helicase motif which is followed by 2 RNase III catalytic domains and a dsRNA-binding motif, similar to the drosha RNase III family. There are close homologs in S. pombe (Acc. Q09884), A. thaliana (Acc. AF187317), D. melanogaster (Acc. AE003740), and human (Acc. AB028449) (Filippov et al., 2000; Jacobsen et al., 1999; Matsuda et al., 2000). Possibly the K12H4.8 RNase III/helicase is the likely candidate to be involved in RNAi.

Genetic screens in C. elegans identified rde-1 and rde-4 as essential for activation of RNAi without an effect on transposon mobilization or co-suppression (Dernburg et al., 2000; Grishok et al., 2000; Ketting and Plasterk, 2000; Tabara et al., 1999). This led to the hypothesis that these genes are important for dsRNA processing but are not involved in mRNA target degradation. The function of both genes is as yet unknown, the rde-1 gene product is a member of a family of proteins similar to the rabbit protein eIF2C (Tabara et al., 1999), and the sequence of rde-4 has not yet

been described. Future biochemical characterization of these proteins should reveal their molecular function.

Processing to the siRNA duplexes appears to start from the ends of both blunt-ended dsRNAs or dsRNAs with short (1-5 nt) 3' overhangs, and proceeds in approximately 21-23 nt steps. Long (~20 nt) 3' staggered ends on short dsRNAs suppress RNAi, possibly through interaction with single-stranded RNA-binding proteins. The suppression of RNAi by single-stranded regions flanking short dsRNA and the lack of siRNA formation from short 30 bp dsRNAs may explain why structured regions frequently encountered in mRNAs do not lead to activation of RNAi.

Without wishing to be bound by theory, we presume that the dsRNA-processing proteins or a subset of these remain associated with the siRNA duplex after the processing reaction. The orientation of the siRNA duplex relative to these proteins determines which of the two complementary strands functions in guiding target RNA degradation. Chemically synthesized siRNA duplexes guide cleavage of sense as well as antisense target RNA as they are able to associate with the protein components in either of the two possible orientation.

The remarkable finding that synthetic 21 and 22 nt siRNA duplexes can be used for efficient mRNA degradation provides new tools for sequence-specific regulation of gene expression in functional genomics as well as biomedical studies. The siRNAs may be effective in mammalian systems where long dsRNAs cannot be used due to the activation of the PKR response (Clemens, 1997). As such, the siRNA duplexes represent a new alternative to antisense or ribozyme therapeutics.

Example 2

RNA Interference in Human Tissue Cultures

2.1 Methods

2.1.1 RNA preparation

21 nt RNAs were chemically synthesized using Expedite RNA phosphoramidites and thymidine phosphoramidite (Proligo, Germany). Synthetic oligonucleotides were deprotected and gel-purified (Example 1), followed by Sep-Pak C18 cartridge (Waters, Milford, MA, USA) purification (Tuschl, 1993). The siRNA sequences targeting GL2 (Acc. X65324) and GL3 luciferase (Acc. U47296) corresponded to the coding regions 153-173 relative to the first nucleotide of the start codon, siRNAs targeting RL (Acc. AF025846) corresponded to region 119-129 after the start codon. Longer RNAs were transcribed with T7 RNA polymerase from PCR products, followed by gel and Sep-Pak purification. The 49 and 484 bp GL2 or GL3 dsRNAs corresponded to position 113-161 and 113-596, respectively, relative to the start of translation; the 50 and 501 bp RL dsRNAs corresponded to position 118-167 and 118-618, respectively. PCR templates for dsRNA synthesis targeting humanized GFP (hG) were amplified from pAD3 (Kehlenbach, 1998), whereby 50 and 501 bp hG dsRNA corresponded to position 118-167 and 118-618, respectively, to the start codon.

For annealing of siRNAs, 20 μM single strands were incubated in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 min at 90°C followed by 1 h at 37°C. The 37°C incubation step was extended overnight for the 50 and 500 bp dsRNAs and these annealing reactions were performed at 8.4 μM and 0.84 μM strand concentrations, respectively.

2.1.2 Cell Culture

S2 cells were propagated in Schneider's *Drosophila* medium (Life Technologies) supplemented with 10% FBS, 100 units/ml penicillin and 100 µg/ml streptomycin at 25°C. 293, NIH/3T3, HeLa S3, COS-7 cells were grown at 37°C in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 100 units/ml penicillin and 100 µg/ml streptomycin. Cells were regularly passaged to maintain exponential growth. 24 h before transfection at approx. 80% confluency, mammalian cells were trypsinized and diluted 1:5 with fresh medium without antibiotics (1-3 x 10^5 cells/ml) and transferred to 24-well plates (500 µl/well). S2 cells were not trypsinized before splitting. Transfection was carried out with Lipofectamine 2000 reagent (Life Technologies) as described by the manufacturer for adherent cell lines. Per well, 1.0 µg pGL2-Control (Promega) or pGL3-Control (Promega), 0.1 µg pRL-TK (Promega) and 0.28 µg siRNA duplex or dsRNA, formulated into liposomes, were applied; the final volume was 600 µl per well. Cells were incubated 20 h after transfection and appeared healthy thereafter. Luciferase expression was subsequently monitored with the Dual luciferase assay (Promega). Transfection efficiencies were determined by fluorescence microscopy for mammalian cell lines after co-transfection of 1.1 µg hGFP-encoding pAD3 and 0.28 µg invGL2 inGL2 siRNA and were 70-90%. Reporter plasmids were amplified in XL-1 Blue (Stratagene) and purified using the Qiagen EndoFree Maxi Plasmid Kit.

2.2 Results and Discussion

To test whether siRNAs are also capable of mediating RNAi in tissue culture, we synthesized 21 nt siRNA duplexes with symmetric 2 nt 3' overhangs directed against reporter genes coding for sea pansy (*Renilla reniformis*) and two sequence variants of firefly (*Photinus pyralis*, GL2 and GL3) luciferases (Fig. 8a, b). The siRNA duplexes were co-transfected with the reporter plasmid combinations pGL2/pRL or pGL3/pRL into *D. melanogaster Schneider S2* cells or mammalian cells using cationic liposomes. Luciferase activities were determined 20 h after transfection. In all cell lines tested,

we observed specific reduction of the expression of the reporter genes in the presence of cognate siRNA duplexes (Fig. 9a-j). Remarkably, the absolute luciferase expression levels were unaffected by non-cognate siRNAs, indicating the absence of harmful side effects by 21 nt RNA duplexes (e.g. Fig. 10a-d for HeLa cells). In *D. melanogaster* S2 cells (Fig. 9a, b), the specific inhibition of luciferases was complete. In mammalian cells, where the reporter genes were 50- to 100-fold stronger expressed, the specific suppression was less complete (Fig. 9c-j). GL2 expression was reduced 3- to 12-fold, GL3 expression 9- to 25-fold and RL expression 1- to 3-fold, in response to the cognate siRNAs. For 293 cells, targeting of RL luciferase by RL siRNAs was ineffective, although GL2 and GL3 targets responded specifically (Fig. 9i, j). The lack of reduction of RL expression in 293 cells may be due to its 5- to 20-fold higher expression compared to any other mammalian cell line tested and/or to limited accessibility of the target sequence due to RNA secondary structure or associated proteins. Nevertheless, specific targeting of GL2 and GL3 luciferase by the cognate siRNA duplexes indicated that RNAi is also functioning in 293 cells.

The 2 nt 3' overhang in all siRNA duplexes, except for uGL2, was composed of (2'-deoxy) thymidine. Substituion of uridine by thymidine in the 3' overhang was well tolerated in the *D. melanogaster* in vitro sytem and the sequence of the overhang was uncritical for target recognition. The thymidine overhang was chosen, because it is supposed to enhance nuclease resistance of siRNAs in the tissue culture medium and within transfected cells. Indeed, the thymidine-modified GL2 siRNA was slightly more potent than the unmodified uGL2 siRNA in all cell lines tested (Fig. 9a, c, e, g, i). It is conceivable that further modifications of the 3' overhanging nucleotides may provide additional benefits to the delivery and stability of siRNA duplexes.

In co-transfection experiments, 25 nM siRNA duplexes with respect to the final volume of tissue culture medium were used (Fig. 9, 10). Increasing

the siRNA concentration to 100 nM did not enhance the specific silencing effects, but started to affect transfection efficiencies due to competition for liposome encapsulation between plasmid DNA and siRNA (data not shown). Decreasing the siRNA concentration to 1.5 nM did not reduce the specific silencing effect (data not shown), even though the siRNAs were now only 2- to 20-fold more concentrated than the DNA plasmids. This indicates that siRNAs are extraordinarily powerful reagents for mediating gene silencing and that siRNAs are effective at concentrations that are several orders of magnitude below the concentrations applied in conventional antisense or ribozyme gene targeting experiments.

In order to monitor the effect of longer dsRNAs on mammalian cells, 50 and 500 bp dsRNAs cognate to the reporter genes were prepared. As non-specific control, dsRNAs from humanized GFP (hG) (Kehlenbach, 1998) was used. When dsRNAs were co-transfected, in identical amounts (not concentrations) to the siRNA duplexes, the reporter gene expression was strongly and unspecifically reduced. This effect is illustrated for HeLa cells as a representative example (Fig. 10a-d). The absolute luciferase activities were decreased unspecifically 10- to 20-fold by 50 bp dsRNA and 20- to 200-fold by 500 bp dsRNA co-transfection, respectively. Similar unspecific effects were observed for COS-7 and NIH/3T3 cells. For 293 cells, a 10- to 20-fold unspecific reduction was observed only for 500 bp dsRNAs. Unspecific reduction in reporter gene expression by dsRNA > 30 bp was expected as part of the interferon response.

Surprisingly, despite the strong unspecific decrease in reporter gene expression, we reproducibly detected additional sequence-specific, dsRNA-mediated silencing. The specific silencing effects, however, were only apparent when the relative reporter gene activities were normalized to the hG dsRNA controls (Fig. 10e, f). A 2- to 10-fold specific reduction in response to cognate dsRNA was observed, also in the other three mammalian cell lines tested (data not shown). Specific silencing effects with

dsRNAs (356-1662 bp) were previously reported in CHO-K1 cells, but the amounts of dsRNA required to detect a 2- to 4-fold specific reduction were about 20-fold higher than in our experiments (Ui-Tei, 2000). Also CHO-K1 cells appear to be deficient in the interferon response. In another report, 293, NIH/3T3 and BHK-21 cells were tested for RNAi using luciferase/lacZ reporter combinations and 829 bp specific lacZ or 717 bp unspecific GFP dsRNA (Caplen, 2000). The failure of detecting RNAi in this case may be due to the less sensitive luciferase/lacZ reporter assay and the length differences of target and control dsRNA. Taken together, our results indicate that RNAi is active in mammalian cells, but that the silencing effect is difficult to detect, if the interferon system is activated by dsRNA > 30 bp.

In summary, we have demonstrated for the first time siRNA-mediated gene silencing in mammalian cells. The use of short siRNAs holds great promise for inactivation of gene function in human tissue culture and the development of gene-specific therapeutics.

Example 3

Specific Inhibition of Gene Expression by RNA Interference

3.1 Materials and Methods

3.1.1 RNA preparation and RNAi assay

Chemical RNA synthesis, annealing, and luciferase-based RNAi assays were performed as described in Examples 1 or 2 or in previous publications (Tuschl et al., 1999; Zamore et al., 2000). All siRNA duplexes were directed against firefly luciferase, and the luciferase mRNA sequence was derived from pGEM-luc (GenBank acc. X65316) as described (Tuschl et al., 1999). The siRNA duplexes were incubated in D. melanogaster RNAi/translation reaction for 15 min prior to addition of mRNAs. Translation-based RNAi assays were performed at least in triplicates.

For mapping of sense target RNA cleavage, a 177-nt transcript was generated, corresponding to the firefly luciferase sequence between positions 113-273 relative to the start codon, followed by the 17-nt complement of the SP6 promoter sequence. For mapping of antisense target RNA cleavage, a 166-nt transcript was produced from a template, which was amplified from plasmid sequence by PCR using 5' primer TAATACGACTCACTATAGAGCCCATATCGTTTCATA (T7 promoter underlined) and 3' primer AGAGGATGGAACCGCTGG. The target sequence corresponds to the complement of the firefly luciferase sequence between positions 50-215 relative to the start codon. Guanylyl transferase labelling was performed as previously described (Zamore et al., 2000). For mapping of target RNA cleavage, 100 nM siRNA duplex was incubated with 5 to 10 nM target RNA in D. melanogaster embryo lysate under standard conditions (Zamore et al., 2000) for 2 h at 25°C. The reaction was stopped by the addition of 8 volumes of proteinase K buffer (200 mM Tris-HCl pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% w/v sodium dodecyl sulfate). Proteinase K (E.M. Merck, dissolved in water) was added to a final concentration of 0.6 mg/ml. The reactions were then incubated for 15 min at 65°C, extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with 3 volumes of ethanol. Samples were located on 6% sequencing gels. Length standards were generated by partial RNase T1 digestion and partial base hydrolysis of the cap-labelled sense or antisense target RNAs.

3.2 Results

3.2.1 Variation of the 3' overhang in duplexes of 21-nt siRNAs

As described above, 2 or 3 unpaired nucleotides at the 3' end of siRNA duplexes were more efficient in target RNA degradation than the respective blunt-ended duplexes. To perform a more comprehensive analysis of the function of the terminal nucleotides, we synthesized five 21-nt sense siRNAs, each displayed by one nucleotide relative to the target RNA, and eight 21-nt antisense siRNAs, each displaced by one nucleotide relative to

the target (Figure 11A). By combining sense and antisense siRNAs, eight series of siRNA duplexes with synthetic overhanging ends were generated covering a range of 7-nt 3' overhang to 4-nt 5' overhang. The interference of siRNA duplexes was measured using the dual luciferase assay system (Tuschl et al., 1999; Zamore et al., 2000). siRNA duplexes were directed against firefly luciferase mRNA, and sea pansy luciferase mRNA was used as internal control. The luminescence ratio of target to control luciferase activity was determined in the presence of siRNA duplex and was normalized to the ratio observed in the absence of dsRNA. For comparison, the interference ratios of long dsRNAs (39 to 504 pb) are shown in Figure 11B. The interference ratios were determined at concentrations of 5 nM for long dsRNAs (Figure 11A) and at 100 nM for siRNA duplexes (Figure 11C-J). The 100 nM concentrations of siRNAs was chosen, because complete processing of 5 nM 504 bp dsRNA would result in 120 nM total siRNA duplexes.

The ability of 21-nt siRNA duplexes to mediate RNAi is dependent on the number of overhanging nucleotides or base pairs formed. Duplexes with four to six 3' overhanging nucleotides were unable to mediate RNAi (Figure 11C-F), as were duplexes with two or more 5' overhanging nucleotides (Figure 11G-J). The duplexes with 2-nt 3' overhangs were most efficient in mediating RNA interference, though the efficiency of silencing was also sequence-dependent, and up to 12-fold differences were observed for different siRNA duplexes with 2-nt 3' overhangs (compare Figure 11D-H). Duplexes with blunted ends, 1-nt 5' overhang or 1- to 3-nt 3' overhangs were sometimes functional. The small silencing effect observed for the siRNA duplex with 7-nt 3' overhang (Figure 11C) may be due to an antisense effect of the long 3' overhang rather than due to RNAi. Comparison of the efficiency of RNAi between long dsRNAs (Fig. 11B) and the most effective 21-nt siRNA duplexes (Fig. 11E, G, H) indicates that a single siRNA duplex at 100 nM concentration can be as effective as 5 nM 504 bp dsRNA.

3.2.2 Length variation of the sense siRNA paired to an invariant 21-nt antisense siRNA

In order to investigate the effect of length of siRNA on RNAi, we prepared 3 series of siRNA duplexes, combining three 21-nt antisense strands with eight, 18- to 25-nt sense strands. The 3' overhang of the antisense siRNA was fixed to 1, 2, or 3 nt in each siRNA duplex series, while the sense siRNA was varied at its 3' end (Figure 12A). Independent of the lenght of the sense siRNA, we found that duplexes with 2-nt 3' overhang of antisense siRNA (Figure 12C) were more active than those with 1- or 3-nt 3' overhang (Figure 12B, D). In the first series, with 1-nt 3' overhang of antisense siRNA, duplexes with a 21- and 22-nt sense siRNAs, carrying a 1- and 2-nt 3' overhang of sense siRNA, respectively, were most active. Duplexes with 19- to 25-nt sense siRNAs were also able to mediate RNA, but to a lesser extent. Similarly, in the second series, with 2-nt overhang of antisense siRNA, the 21-nt siRNA duplex with 2-nt 3' overhang was most active, and any other combination with the 18- to 25-nt sense siRNAs was active to a significant degree. In the last series, with 3-nt antisense siRNA 3' overhang, only the duplex with a 20-nt sense siRNA and the 2-nt sense 3' overhang was able to reduce target RNA expression. Together, these results indicate that the length of the siRNA as well as the length of the 3' overhang are important, and that duplexes of 21-nt siRNAs with 2-nt 3' overhang are optimal for RNAi.

3.2.3 Length variation of siRNA duplexes with a constant 2-nt 3' overhang

We then examined the effect of simultaneously changing the length of both siRNA strands by maintaining symmetric 2-nt 3' overhangs (Figure 13A). Two series of siRNA duplexes were prepared including the 21-nt siRNA duplex of Figure 11H as reference. The length of the duplexes was varied between 20 to 25 bp by extending the base-paired segment at the 3' end of the sense siRNA (Figure 13B) or at the 3' end of the antisense siRNA (Figure 13C). Duplexes of 20 to 23 bp caused specific repression of target luciferase activity, but the 21-nt siRNA duplex was at least 8-fold more

efficient than any of the other duplexes. 24- and 25-nt siRNA duplexes did not result in any detectable interference. Sequence-specific effects were minor as variations on both ends of the duplex produced similar effects.

3.2.4 2'-Deoxy and 2'-O-methyl-modified siRNA duplexes

To assess the importance of the siRNA ribose residues for RNAi, duplexes with 21-nt siRNAs and 2-nt 3' overhangs with 2'-deoxy- or 2'-O-methyl-modified strands were examined (Figure 14). Substitution of the 2-nt 3' overhangs by 2'-deoxy nucleotides had no effect, and even the replacement of two additional ribonucleotides adjacent to the overhangs in the paired region, produced significantly active siRNAs. Thus, 8 out of 42 nt of a siRNA duplex were replaced by DNA residues without loss of activity. Complete substitution of one or both siRNA strands by 2'-deoxy residues, however, abolished RNAi, as did substitution by 2'-O-methyl residues.

3.2.5 Definition of target RNA cleavage sites

Target RNA cleavage positions were previously determined for 22-nt siRNA duplexes and for a 21-nt/22-nt duplex. It was found that the position of the target RNA cleavage was located in the centre of the region covered by the siRNA duplex, 11 or 12 nt downstream of the first nucleotide that was complementary to the 21- or 22-nt siRNA guide sequence. Five distinct 21-nt siRNA duplexes with 2-nt 3' overhang (Figure 15A) were incubated with 5' cap-labelled sense or antisense target RNA in D. melanogaster lysate (Tuschl et al., 1999; Zamore et al., 2000). The 5' cleavage products were resolved on sequencing gels (Figure 15B). The amount of sense target RNA cleaved correlates with the efficiency of siRNA duplexes determined in the translation-based assay, and siRNA duplexes 1, 2 and 4 (Figure 15B and 11H, G, E) cleave target RNA faster than duplexes 3 and 5 (Figure 15B and 11F, D). Notably, the sum of radioactivity of the 5' cleavage product and the input target RNA were not constant over time, and the 5' cleavage products did not accumulate. Presumably, the cleavage products, once

released from the siRNA-endonuclease complex, are rapidly degraded due to the lack of either of the poly(A) tail of the 5'-cap.

The cleavage sites for both, sense and antisense target RNAs were located in the middle of the region spanned by the siRNA duplexes. The cleavage sites for each target produced by the 5 different duplexes varied by 1-nt according to the 1-nt displacement of the duplexes along the target sequences. The targets were cleaved precisely 11 nt downstream of the target position complementary to the 3'-most nucleotide of the sequence-complementary guide siRNA (Figure 15A, B).

In order to determine, whether the 5' or the 3' end of the guide siRNA sets the ruler for target RNA cleavage, we devised the experimental strategy outlined in Figure 16A and B. A 21-nt antisense siRNA, which was kept invariant for this study, was paired with sense siRNAs that were modified at either of their 5' or 3' ends. The position of sense and antisense target RNA cleavage was determined as described above. Changes in the 3' end of the sense siRNA, monitored for 1-nt 5' overhang to 6-nt 3' overhang, did neither effect the position of sense nor antisense target RNA cleavage (Figure 16C). Changes in the 5' end of the sense siRNA did no affect the sense target RNA cleavage (Figure 16D, top panel), which was expected because the antisense siRNA was unchanged. However, the antisense target RNA cleavage was affected and strongly dependent on the 5' end of the sense siRNA (Figure 16D, bottom panel). The antisense target was only cleaved, when the sense siRNA was 20 or 21 nt in size, and the position of cleavage different by 1-nt, suggesting that the 5' end of the target-recognizing siRNA sets the ruler for target RNA cleavage. The position is located between nucleotide 10 and 11 when counting in upstream direction from the target nucleotide paired to the 5'-most nucleotide of the guide siRNA (see also Figure 15A).

3.2.6 Sequence effects and 2'-deoxy substitutions in the 3' overhang

A 2-nt 3'overhang is preferred for siRNA function. We wanted to know, if the sequence of the overhanging nucleotides contributes to target recognition, or if it is only a feature required for reconstitution of the endonuclease complex (RISC or siRNP). We synthesized sense and antisense siRNAs with AA, CC, GG, UU, and UG 3' overhangs and included the 2'-deoxy modifications TdG and TT. The wild-type siRNAs contained AA in the sense 3' overhang and UG in the antisense 3' overhang (AA/UG). All siRNA duplexes were functional in the interference assay and reduced target expression at least 5-fold (Figure 17). The most efficient siRNA duplexes that reduced target expression more than 10-fold, were of the sequence type NN/UG, NN/UU, NN/TdG, and NN/TT (N, any nucleotide). siRNA duplexes with an antisense siRNA 3' overhang of AA, CC or GG were less active by a factor 2 to 4 when compared to the wild-type sequence UG or the mutant UU. This reduction in RNAi efficiency is likely due to the contribution of the penultimate 3' nucleotide to sequence-specific target recognition, as the 3' terminal nucleotide was changed from G to U without effect.

Changes in the sequence of the 3' overhang of the sense siRNA did not reveal any sequence-dependent effects, which was expected, because the sense siRNA must not contribute to sense target mRNA recognition.

3.2.7 Sequence specifity of target recognition

In order to examine the sequence-specifity of target recognition, we introduced sequence changes into the paired segments of siRNA duplexes and determined the efficiency of silencing. Sequence changes were introduced by inverting short segments of 3- or 4-nt length or as point mutations (Figure 18). The sequence changes in one siRNA strand were compensated in the complementary siRNA strand to avoid pertubing the base-paired siRNA duplex structure. The sequence of all 2-nt 3' overhangs was TT (T, 2'-deoxythymidine) to reduce costs of synthesis. The TT/TT refe-

rence siRNA duplex was comparable in RNAi to the wild-type siRNA duplex AA/UG (Figure 17). The ability to mediate reporter mRNA destruction was quantified using the translation-based luminescence assay. Duplexes of siRNAs with inverted sequence segments showed dramatically reduced ability for targeting the firefly luciferase reporter (Figure 18). The sequence changes located between the 3' end and the middle of the antisense siRNA completely abolished target RNA recognition, but mutations near the 5' end of the antisense siRNA exhibit a small degree of silencing. Transversion of the A/U base pair located directly opposite of the predicted target RNA cleavage site, or one nucleotide further away from the predicted site, prevented target RNA cleavage, therefore indicating that single mutation within the centre of a siRNA duplex discriminate between mismatched targets.

3.3 Discussion

siRNAs are valuable reagents for inactivation of gene expression, not only in insect cells, but also in mammalian cells, with a great potential for therapeutic application. We have systematically analysed the structural determinants of siRNA duplexes required to promote efficient target RNA degradation in D. melanogaster embryo lysate, thus providing rules for the design of most potent siRNA duplexes. A perfect siRNA duplex is able to silence gene expression with an efficiency comparable to a 500 bp dsRNA, given that comparable quantities of total RNA are used.

3.4 The siRNA user guide

Efficiently silencing siRNA duplexes are preferably composed of 21-nt antisense siRNAs, and should be selected to form a 19 bp double helix with 2-nt 3' overhanging ends. 2'-deoxy substitutions of the 2-nt 3' overhanging ribonucleotides do not affect RNAi, but help to reduce the costs of RNA synthesis and may enhance RNAse resistance of siRNA duplexes. More extensive 2'-deoxy or 2'-O-methyl modifications, however, reduce

the ability of siRNAs to mediate RNAi, probably by interfering with protein association for siRNAP assembly.

Target recognition is a highly sequence-specific process, mediated by the siRNA complementary to the target. The 3'-most nucleotide of the guide siRNA does not contribute to specificity of target recognition, while the penultimate nucleotide of the 3' overhang affects target RNA cleavage, and a mismatch reduces RNAi 2- to 4-fold. The 5' end of a guide siRNA also appears more permissive for mismatched target RNA recognition when compared to the 3' end. Nucleotides in the centre of the siRNA, located opposite the target RNA cleavage site, are important specificity determinants and even single nucleotide changes reduce RNAi to undetectable level. This suggests that siRNA duplexes may be able to discriminate mutant or polymorphic alleles in gene targeting experiments, which may become an important feature for future therapeutic developments.

Sense and antisense siRNAs, when associated with the protein components of the endonclease complex or its commitment complex, were suggested to play distinct roles; the relative orientation of the siRNA duplex in this complex defines which strand can be used for target recognition. Synthetic siRNA duplexes have dyad symmetry with respect to the double-helical structure, but not with respect to sequence. The association of siRNA duplexes with the RNAi proteins in the D. melanogaster lysate will lead to formation of two asymmetric complexes. In such hypothetical complexes, the chiral environment is distinct for sense and antisense siRNA, hence their function. The prediction obviously does not apply to palindromic siRNA sequences, or to RNAi proteins that could associate as homodimers. To minimize sequence effects, which may affect the ratio of sense and antisense-targeting siRNPs, we suggest to use siRNA sequences with identical 3' overhanging sequences. We recommend to adjust the sequence of the overhang of the sense siRNA to that of the antisense 3' overhang, because the sense siRNA does not have a target in typical

knock-down experiments. Asymmetry in reconstitution of sense and anti-sense-cleaving siRNPs could be (partially) responsible for the variation in RNAi efficiency observed for various 21-nt siRNA duplexes with 2-nt 3' overhangs used in this study (Figure 14). Alternatively, the nucleotide sequence at the target site and/or the accessibility of the target RNA structure may be responsible for the variation in efficiency for these siRNA duplexes.

References

Bass, B. L. (2000). Double-stranded RNA as a template for gene silencing. Cell 101, 235-238.

Bosher, J. M., and Labouesse, M. (2000). RNA interference: genetic wand and genetic watchdog. Nat. Cell Biol. 2, E31-36.

Caplen, N. J., Fleenor, J., Fire, A., and Morgan, R. A. (2000). dsRNA-mediated gene silencing in cultured Drosophila cells: a tissue culture model for the analysis of RNA interference. Gene 252, 95-105.

Catalanotto, C., Azzalin, G., Macino, G., and Cogoni, C. (2000). Gene silencing in worms and fungi. Nature 404, 245.

Chanfreau, G., Buckle, M., and Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc. Natl. Acad. Sci. USA 97, 3142-3147.

Clemens, M. J. (1997). PKR--a protein kinase regulated by double-stranded RNA. Int. J. Biochem. Cell Biol. 29, 945-949.

Cogoni, C., and Macino, G. (1999). Homology-dependent gene silencing in plants and fungi: a number of variations on the same theme. Curr. Opin. Microbiol. 2, 657-662.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D. C. (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell 101, 543-553.

Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the C. elegans germ line. Genes & Dev. 14, 1578-1583.

Dunn, J. J. (1982). Ribonuclease III. In The enzymes, vol 15, part B, P. D. Boyer, ed. (New York: Academic Press), pp. 485-499.

Filippov, V., Solovyev, V., Filippova, M., and Gill, S. S. (2000). A novel type of RNase III family proteins in eukaryotes. Gene 245, 213-221.

Fire, A. (1999). RNA-triggered gene silencing. Trends Genet. 15, 358-363.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806-811.

Grishok, A., Tabara, H., and Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science 287, 2494-2497.

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286, 950-952.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404, 293-296.

Jacobsen, S. E., Running, M. P., and M., M. E. (1999). Disruption of an RNA helicase/RNase III gene in Arabidopsis causes unregulated cell division in floral meristems. Development 126, 5231-5243.

Jensen, S., Gassama, M. P., and Heidmann, T. (1999). Taming of transposable elements by homology-dependent gene silencing. Nat. Genet. 21, 209-212.

Kehlenbach, R. H., Dickmanns, A. & Gerace, L. (1998). Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT In vitro. J. Cell Biol. 141, 863-874.

Kennerdell, J. R., and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell 95, 1017-1026.

Ketting, R. F., Haverkamp, T. H., van Luenen, H. G., and Plasterk, R. H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell 99, 133-141.

Ketting, R. F., and Plasterk, R. H. (2000). A genetic link between co-suppression and RNA interference in C. elegans. Nature 404, 296-298.

Lucy, A. P., Guo, H. S., Li, W. X., and Ding, S. W. (2000). Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus. EMBO J. 19, 1672-1680.

Matsuda, S., Ichigotani, Y., Okuda, T., Irimura, T., Nakatsugawa, S., and Hamaguchi, M. (2000). Molecular cloning and characterization of a novel human gene (HERNA) which encodes a putative RNA-helicase. Biochim. Biophys. Acta 31, 1-2.

Milligan, J.F., and Uhlenbeck, O.C. (1989). Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol. 180, 51-62.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Jouette, D., Lacombe, A. M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T. A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101, 533-542.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proc. Natl. Acad. Sci. USA 95, 14687-14692.

Nicholson, A. W. (1999). Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol. Rev. 23, 371-390.

Oelgeschlager, M., Larrain, J., Geissert, D., and De Robertis, E. M. (2000). The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling. Nature 405, 757-763.

Pan, T., and Uhlenbeck, O. C. (1992). In vitro selection of RNAs that undergo autolytic cleavage with Pb^{2+}. Biochemistry 31, 3887-3895.

Pelissier, T., and Wassenegger, M. (2000). A DNA target of 30 bp is sufficient for RNA-directed methylation. RNA 6, 55-65.

Plasterk, R. H., and Ketting, R. F. (2000). The silence of the genes. Curr. Opin. Genet. Dev. 10, 562-567.

Ratcliff, F. G., MacFarlane, S. A., and Baulcombe, D. C. (1999). Gene Silencing without DNA. RNA-mediated cross-protection between viruses. Plant Cell 11, 1207-1216.

Robertson, H. D. (1990). Escherichia coli ribonuclease III. Methods Enzymol. 181, 189-202.

Robertson, H. D. (1982). Escherichia coli ribonuclease III cleavage sites. Cell 30, 669-672.

Romaniuk, E., McLaughlin, L. W., Neilson, T., and Romaniuk, P. J. (1982). The effect of acceptor oligoribonucleotide sequence on the T4 RNA ligase reaction. Eur J Biochem 125, 639-643.

Sharp, P. A. (1999). RNAi and double-strand RNA. Genes & Dev. 13, 139-141.

Sijen, T., and Kooter, J. M. (2000). Post-transcriptional gene-silencing: RNAs on the attack or on the defense? Bioessays 22, 520-531.

Smardon, A., Spoerke, J., Stacey, S., Klein, M., Mackin, N., and Maine, E. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr. Biol. 10, 169-178.

Svoboda, P., Stein, P., Hayashi, H., and Schultz, R. M. (2000). Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. Development 127, 4147-4156.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99, 123-132.

Tuschl, T., Ng, M. M., Pieken, W., Benseler, F., and Eckstein, F. (1993). Importance of exocyclic base functional groups of central core guanosines for hammerhead ribozyme activity. Biochemistry 32, 11658-11668.

Tuschl, T., Sharp, P. A., and Bartel, D. P. (1998). Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. EMBO J. 17, 2637-2650.

Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. Genes & Dev. 13, 3191-3197.

Ui-Tei, K., Zenno, S., Miyata, Y. & Saigo, K. (2000). Sensitive assay of RNA interference in Drosophila and Chinese hamster cultured cells using firefly luciferase gene as target. FEBS Letters 479, 79-82.

Verma, S., and Eckstein, F. (1999). Modified oligonucleotides: Synthesis and strategy for users. Annu. Rev. Biochem. 67, 99-134.

Voinnet, O., Lederer, C., and Baulcombe, D. C. (2000). A viral movement protein prevents spread of the gene silencing signal in Nicotiana benthamiana. Cell 103, 157-167.

Wassenegger, M. (2000). RNA-directed DNA methylation. Plant Mol. Biol. 43, 203-220.

Wianny, F., and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. Nat. Cell Biol. 2, 70-75.

Wu, H., Xu, H., Miraglia, L. J., and Crooke, S. T. (2000). Human RNase III is a 160 kDa Protein Involved in Preribosomal RNA Processing. J. Biol. Chem. 17, 17.

Yang, D., Lu, H. and Erickson, J.W. (2000) Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in drosophilia embryos. Curr. Biol., 10, 1191-1200.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101, 25-33.

Zhang, K., and Nicholson, A. W. (1997). Regulation of ribonuclease III processing by double-helical sequence antideterminants. Proc. Natl. Acad. Sci. USA 94, 13437-13441.

Claims

1. Isolated double-stranded RNA molecule, wherein each RNA strand has a length from 19-25 nucleotides, wherein said RNA molecule is capable of target-specific nucleic acid modifications.

2. The RNA molecule of claim 1 wherein at least one strand has a 3'-overhang from 1-5 nucleotides.

3. The RNA molecule of claim 1 or 2 capable of target-specific RNA interference and/or DNA methylation.

4. The RNA molecule of any one of claims 1-3, wherein each strand has a length from 19-23, particularly from 20-22 nucleotides.

5. The RNA molecule of any one of claims 2-4, wherein the 3'-overhang is from 1-3 nucleotides.

6. The RNA molecule of any one of claims 2-5, wherein the 3'-overhang is stabilized against degradation.

7. The RNA molecule of any one of claims 1-6, which contains at least one modified nucleotide analogue.

8. The RNA molecule of claim 7, wherein the modified nucleotide analogue is selected from sugar- or backbone modified ribonucleotides.

9. The RNA molecule according to claim 7 or 8, wherein the nucleotide analogue is a sugar-modified ribonucleotide, wherein the 2'-OH group is replaced by a group selected from H, OR, R, halo, SH, SR^1,

NH_2, NHR, NR_2 or CN, wherein R is C_1-C_6 alkyl, alkenyl or alkynyl and halo is F, Cl, Br or I.

10. The RNA molecule of claim 7 or 8, wherein the nucleotide analogue is a backbone-modified ribonucleotide containing a phosphothioate group.

11. The RNA molecule of any one of claims 1-10, which has a sequence having an identity of at least 50 percent to a predetermined mRNA target molecule.

12. The RNA molecule of claim 11, wherein the identity is at least 70 percent.

13. A method of preparing a double-stranded RNA molecule of any one of claims 1-12 comprising the steps:

(a) synthesizing two RNA strands each having a length from 19-25 nucleotides, wherein said RNA strands are capable of forming a double-stranded RNA molecule,

(b) combining the synthesized RNA strands under conditions, wherein a double-stranded RNA molecule is formed, which is capable of target-specific nucleic acid modifications.

14. The method of claim 13, wherein the RNA strands are chemically synthesized.

15. The method of claim 13, wherein the RNA strands are enzymatically synthesized.

16. A method of mediating target-specific nucleic acid modifications in a cell or an organism comprising the steps:

(a) contacting said cell or organism with the double-stranded RNA molecule of any one of claims 1-12 under conditions wherein target-specific nucleic acid modifications can occur, and

(b) mediating a target-specific nucleic acid modification effected by the double-stranded RNA towards a target nucleic acid having a sequence portion substantially corresponding to the double-stranded RNA.

17. The method of claim 16, wherein the nucleic acid modification is RNA interference and/or DNA methylation.

18. The method of claim 16 and 17 wherein said contacting comprises introducing said double-stranded RNA molecule into a target cell in which the target-specific nucleic acid modification can occur.

19. The method of claim 18 wherein the introducing comprises a carrier-mediated delivery or injection.

20. Use of the method of any one of claims 16-19 for determining the function of a gene in a cell or an organism.

21. Use of the method of any one of claims 16-19 for modulating the function of a gene in a cell or an organism.

22. The use of claim 20 or 21, wherein the gene is associated with a pathological condition.

23. The use of claim 22, wherein the gene is a pathogen-associated gene.

24. The use of claim 23, wherein the gene is a viral gene.

25. The use of claim 22, wherein the gene is a tumor-associated gene.

26. The use of claim 22, wherein the gene is an autoimmune disease-associated gene.

27. Pharmaceutical composition containing as an active agent at least one double-stranded RNA molecule of any one of claims 1-12 and a pharmaceutical carrier.

28. The composition of claim 27 for diagnostic applications.

29. The composition of claim 27 for therapeutic applications.

30. A eukaryotic cell or a eukaryotic non-human organism exhibiting a target gene-specific knockout phenotype wherein said cell or organism is transfected with at least one double-stranded RNA molecule capable of inhibiting the expression of an endogeneous target gene or with a DNA encoding at least one double-stranded RNA molecule capable of inhibiting the expression of at least one endogeneous target gene.

31. The cell or organism of claim 30 which is a mammalian cell.

32. The cell or organism of claim 31 which is a human cell.

33. The cell or organism of any one of claims 30-32 which is further transfected with at least one exogeneous target nucleic acid coding

for the target protein or a variant or mutated form of the target protein, wherein said exogeneous target nucleic acid differs from the endogeneous target gene on the nucleic acid level such that the expression of the exogeneous target nucleic acid is substantially less inhibited by the double stranded RNA molecule than the expression of the endogeneous target gene.

34. The cell or organism of claim 33 wherein the exogeneous target nucleic acid is fused to a further nucleic acid sequence encoding a detectable peptide or polypeptide.

35. Use of the cell or organism of any of claims 30-34 for analytic procedures.

36. The use of claim 35 for the analysis of gene expression profiles.

37. The use of claim 35 for a proteome analysis.

38. The use of any one of claims 35-37 wherein an analysis of a variant or mutant form of the target protein encoded by an exogeneous target nucleic acid is carried out.

39. The use of claim 38 for identifying functional domains of the target protein.

40. The use of any one of claims 35-39 wherein a comparison of at least two cells or organisms is carried out selected from:

(i) a control cell or control organism without target gene inhibition,

(ii) a cell or organism with target gene inhibition and

(iii) a cell or organism with target gene inhibition plus target gene complementation by an exogeneous target nucleic acid.

41. The use of any one of claims 35-40 wherein the analysis comprises a functional and/or phenotypic analysis.

42. Use of a cell of any one of claims 30-34 for preparative procedures.

43. The use of claim 41 for the isolation of proteins or protein complexes from eukaryotic cells.

44. The use of claim 43 for the isolation of high molecular weight protein complexes which may optionally contain nucleic acids.

45. The use of any one of claims 35-44 in a procedure for identifying and/or characterizing pharmacological agents.

46. A system for identifying and/or characterizing a pharmacological agent acting on at least one target protein comprising:

(a) a eukaryotic cell or a eukaryotic non-human organism capable of expressing at least one target gene coding for said at least one target protein,

(b) at least one double-stranded RNA molecule capable of inhibiting the expression of said at least one endogeneous target gene, and

(c) a test substance or a collection of test substances wherein pharmacological properties of said test substance or said collection are to be identified and/or characterized.

47. The system of claim 46 further comprising:

(d) at least one exogeneous target nucleic acid coding for the target protein or a variant or mutated from of the target protein wherein said exogeneous target nucleic acid differs from the endogeneous target gene on the nucleic acid level such that the expression of the exogeneous target nucleic

acid is substantially less inhibited by the double stranded RNA molecule than the expression of the endogeneous target gene.

FIGURE 1A

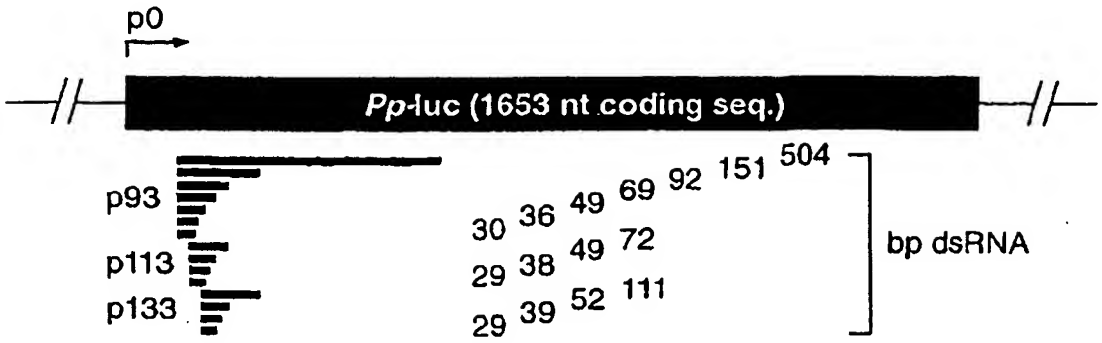


FIGURE 1B

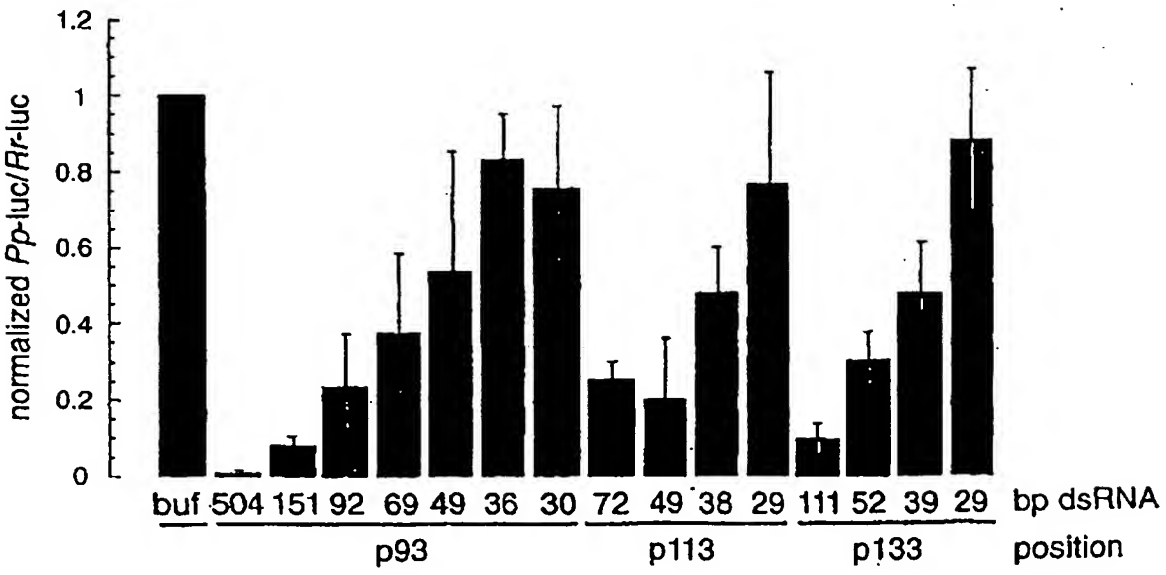


FIGURE 2

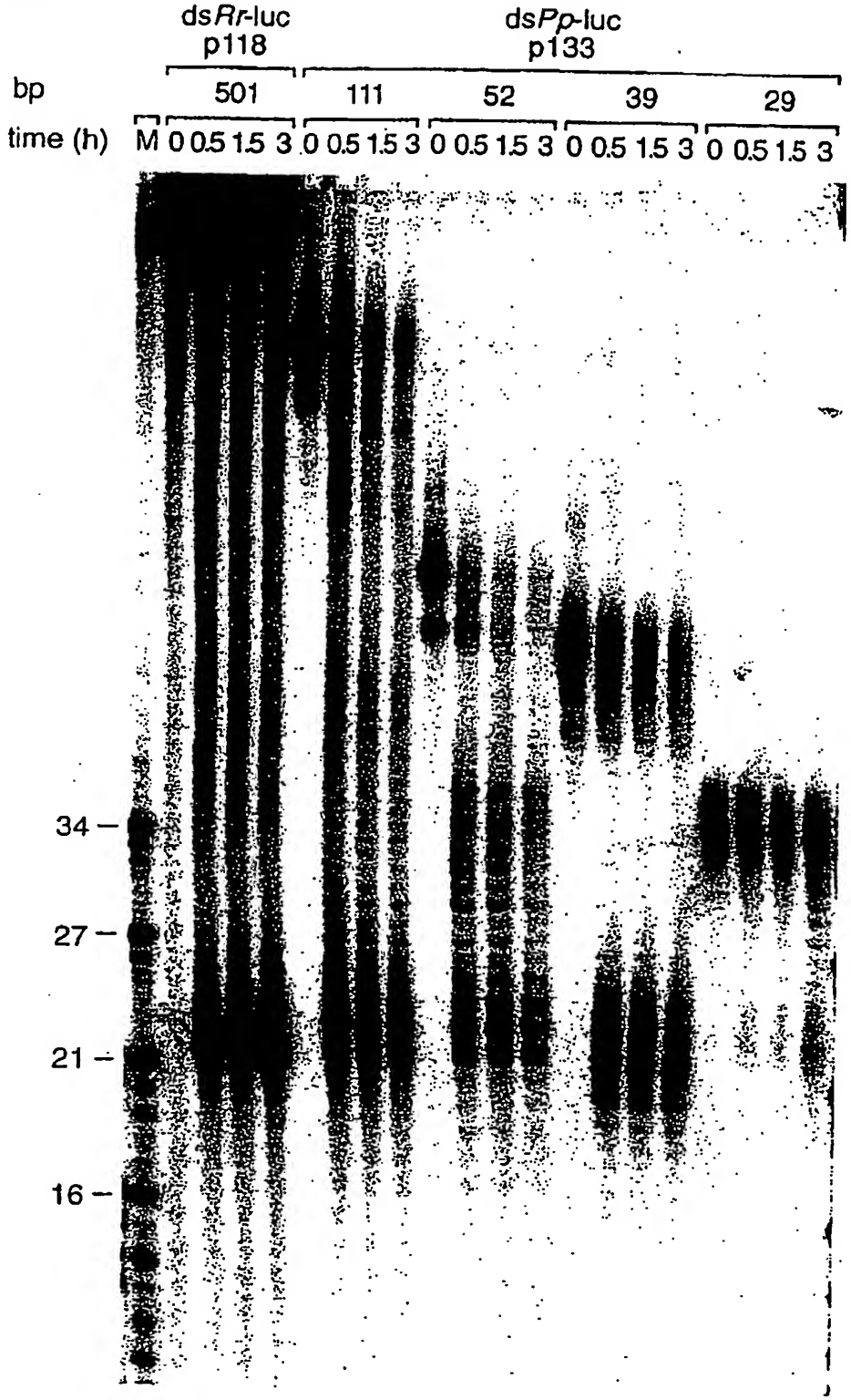

ds*Rr*-luc p118
ds*Pp*-luc p133
bp
501
111
52
39
29
time (h)
M 0 0.5 1.5 3 0 0.5 1.5 3 0 0.5 1.5 3 0 0.5 1.5 3 0 0.5 1.5 3
34
27
21
16

Figure 3A

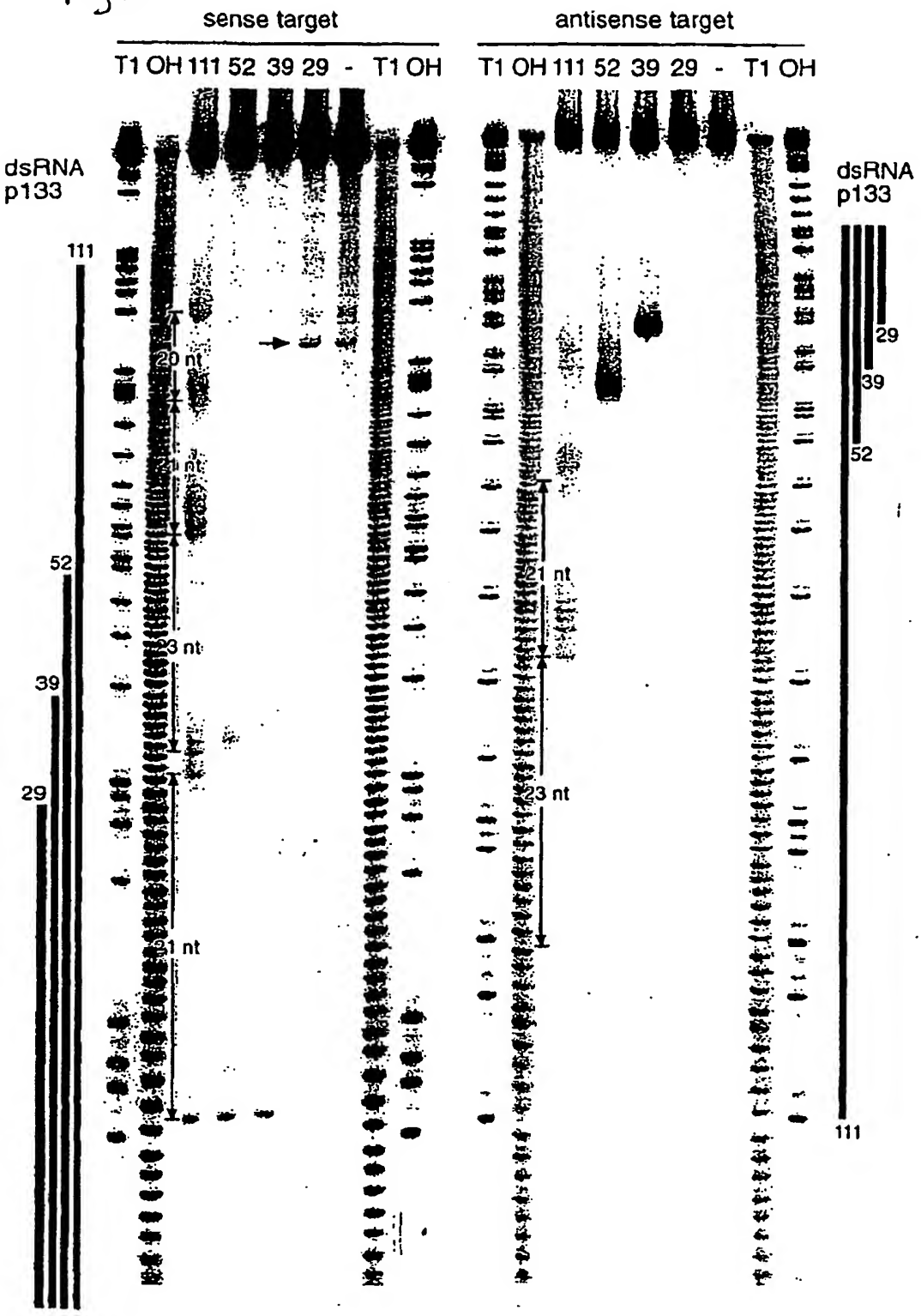

sense target
antisense target
T1 OH 111 52 39 29 - T1 OH
T1 OH 111 52 39 29 - T1 OH
dsRNA p133
dsRNA p133
111
52
39
29
20 nt
1 nt
3 nt
21 nt
1 nt
23 nt
29
39
52
111

FIGURE 3B

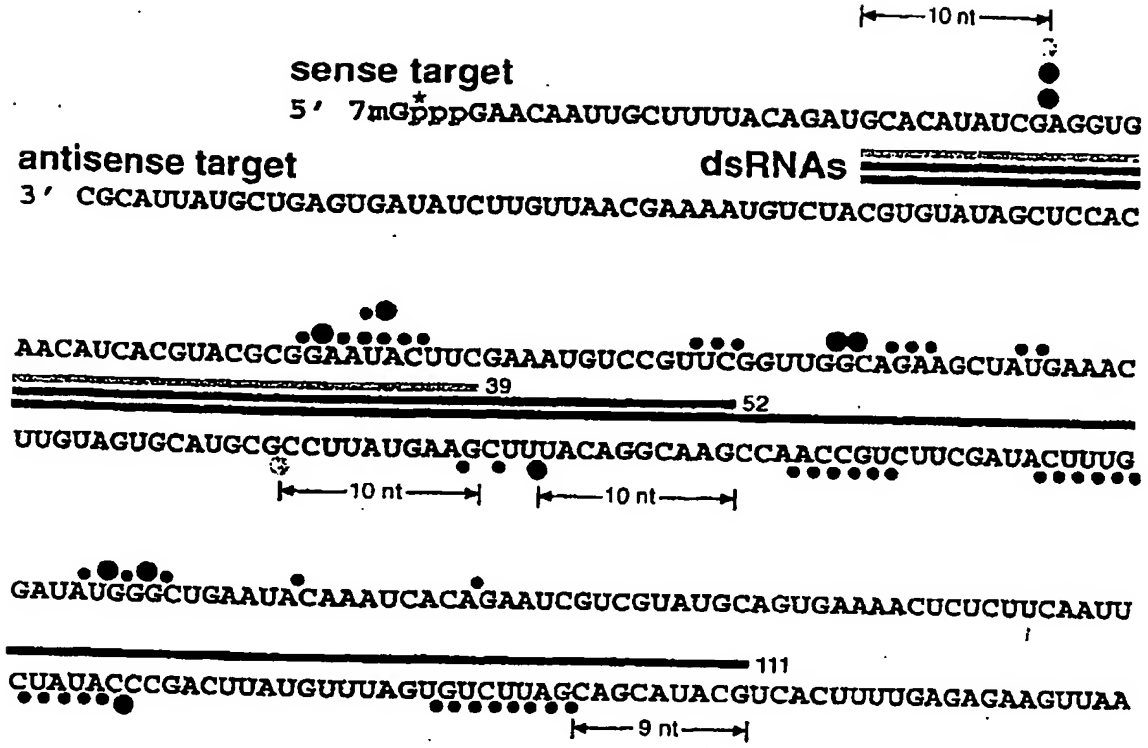
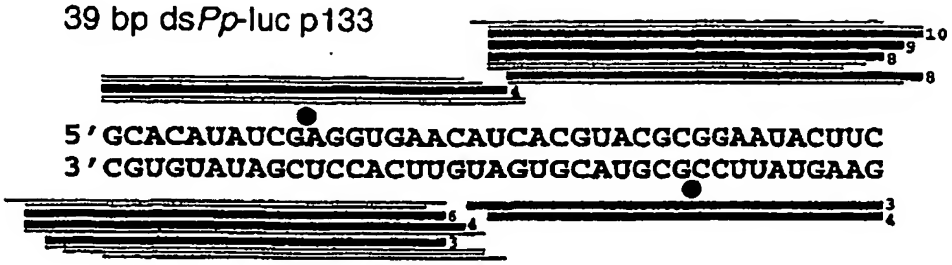
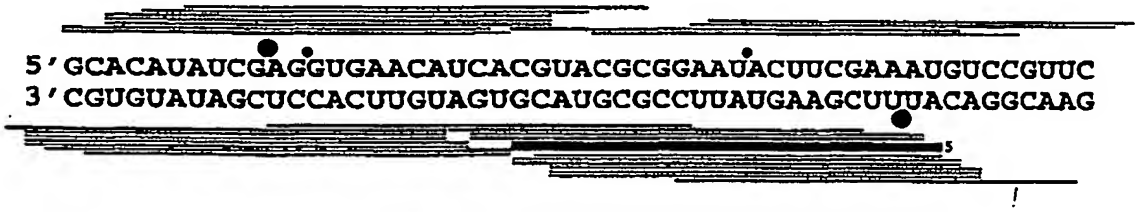

10 nt
sense target
5' 7mGpppGAACAAUUGCUUUUACAGAUGCACAUAUCGAGGUG
antisense target
dsRNAs
3' CGCAUUAUGCUGAGUGAUAUCUUGUUAACGAAAAUGUCUACGUGUAUAGCUCCAC
AACAUCACGUACGCGGAAUACUUCGAAAUGUCCGUUCGGUUGGCAGAAGCUAUGAAAC
39
52
UUGUAGUGCAUGCGCCUUAUGAAGCUUUACAGGCAAGCCAACCGUCUUCGAUACUUUG
10 nt
10 nt
GAUAUGGGCUGAAUACAAAUCACAGAAUCGUCGUAUGCAGUGAAAACUCUCUUCAAUU
111
CUAUACCCGACUUAUGUUUAGUGUCUUAGCAGCAUACGUCACUUUUGAGAGAAGUUAA
9 nt
CUUUAUGCCUAUAGUGUCACCUAAAU 3'
GAAAUACGGpppGm7 5'

FIGURE 4A

39 bp ds*Pp*-luc p133



52 bp ds*Pp*-luc p133



111 bp ds*Pp*-luc p133

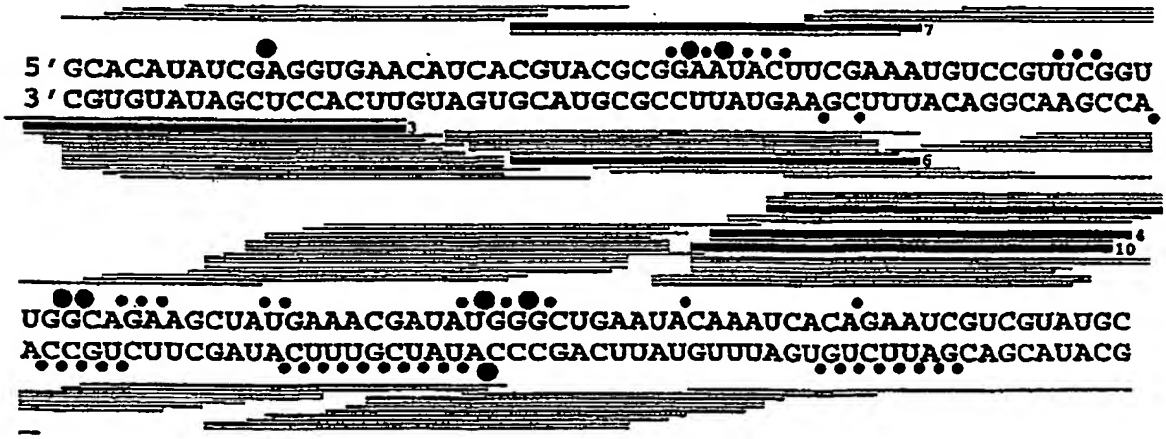


FIGURE 4B

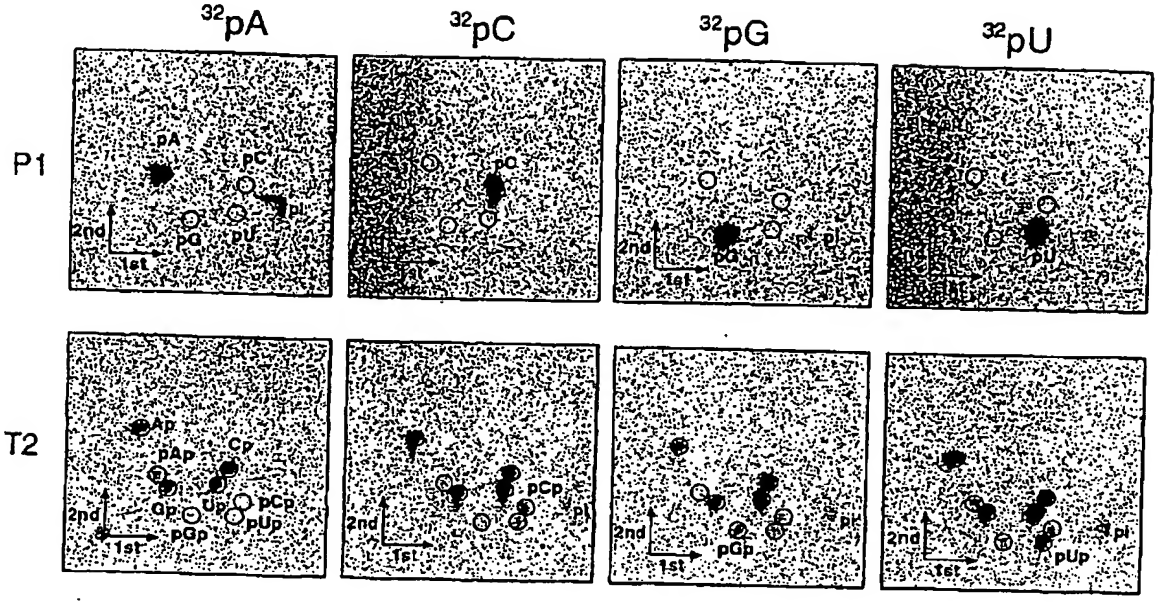

32pA
32pC
32pG
32pU
P1
T2
pA
pC
pG
pU
pi
2nd
1st
Ap
pAp
Cp
Gp
Up
pCp
pUp
pGp

Figure 5A

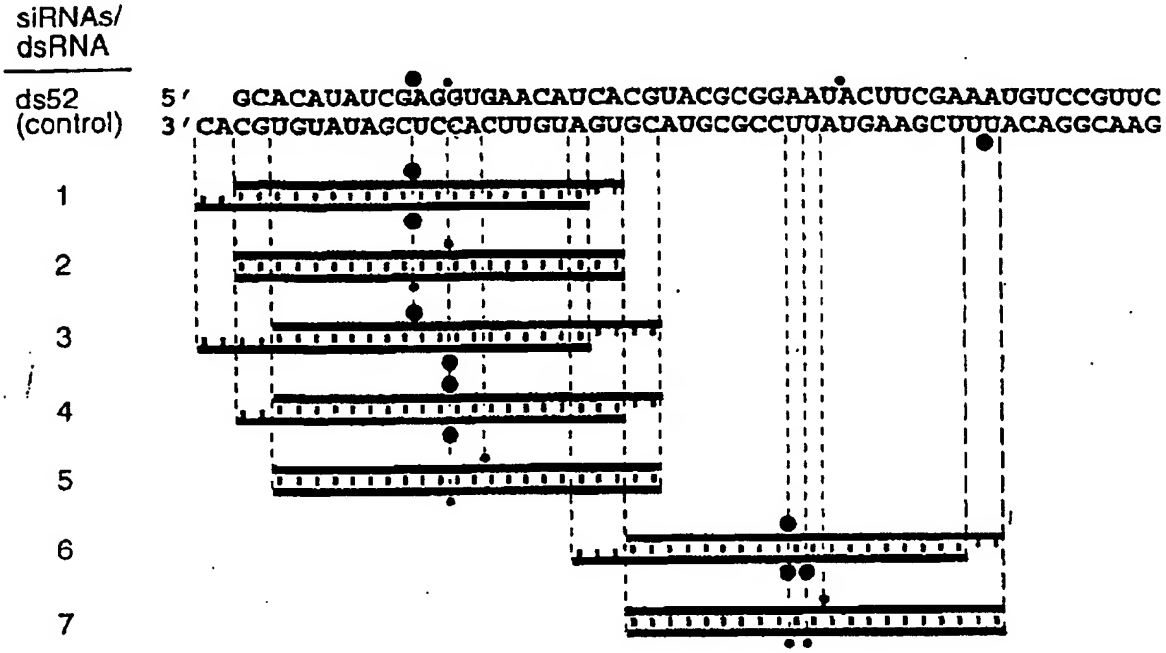

siRNAs/
dsRNA
ds52
(control)
5′ GCACAUAUCGAGGUGAACAUCACGUACGCGGAAUACUUCGAAAUGUCCGUUC
3′ CACGUGUAUAGCUCCACUUGUAGUGCAUGCGCCUUAUGAAGCUUUACAGGCAAG
1
2
3
4
5
6
7

FIGURE 5B

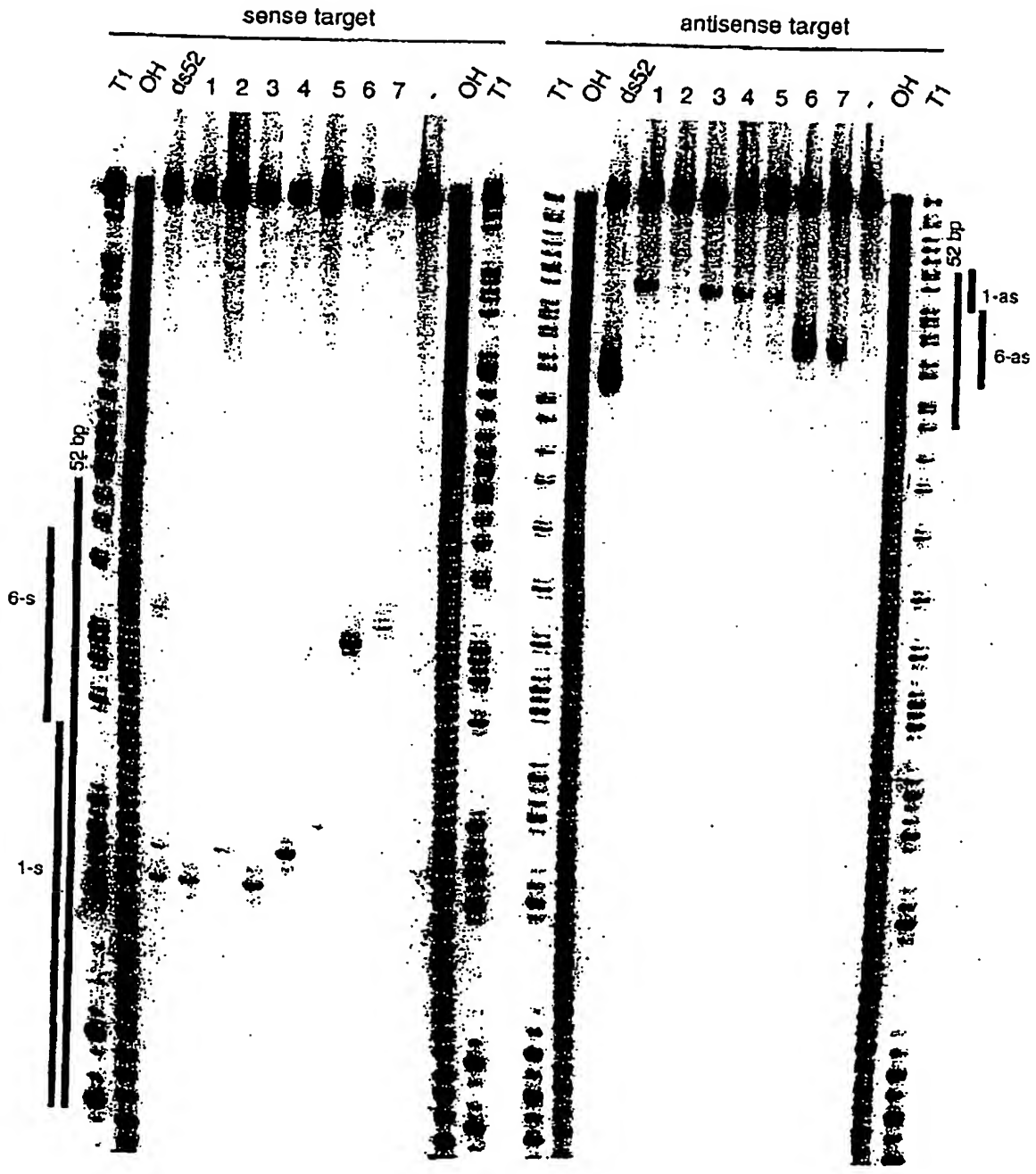

sense target
antisense target
T1
OH
ds52
1
2
3
4
5
6
7
OH
T1
52 bp
6-s
1-s
1-as
6-as

FIGURE 6A

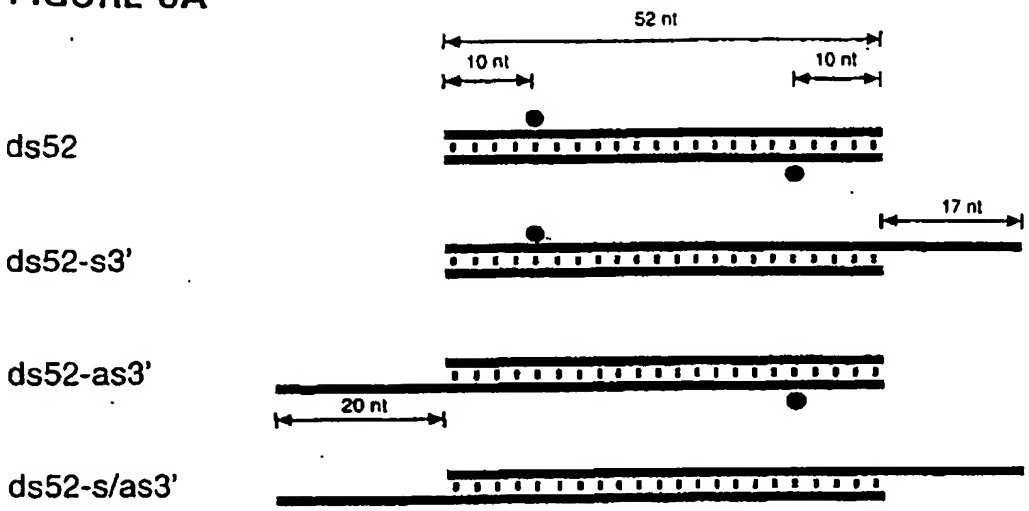
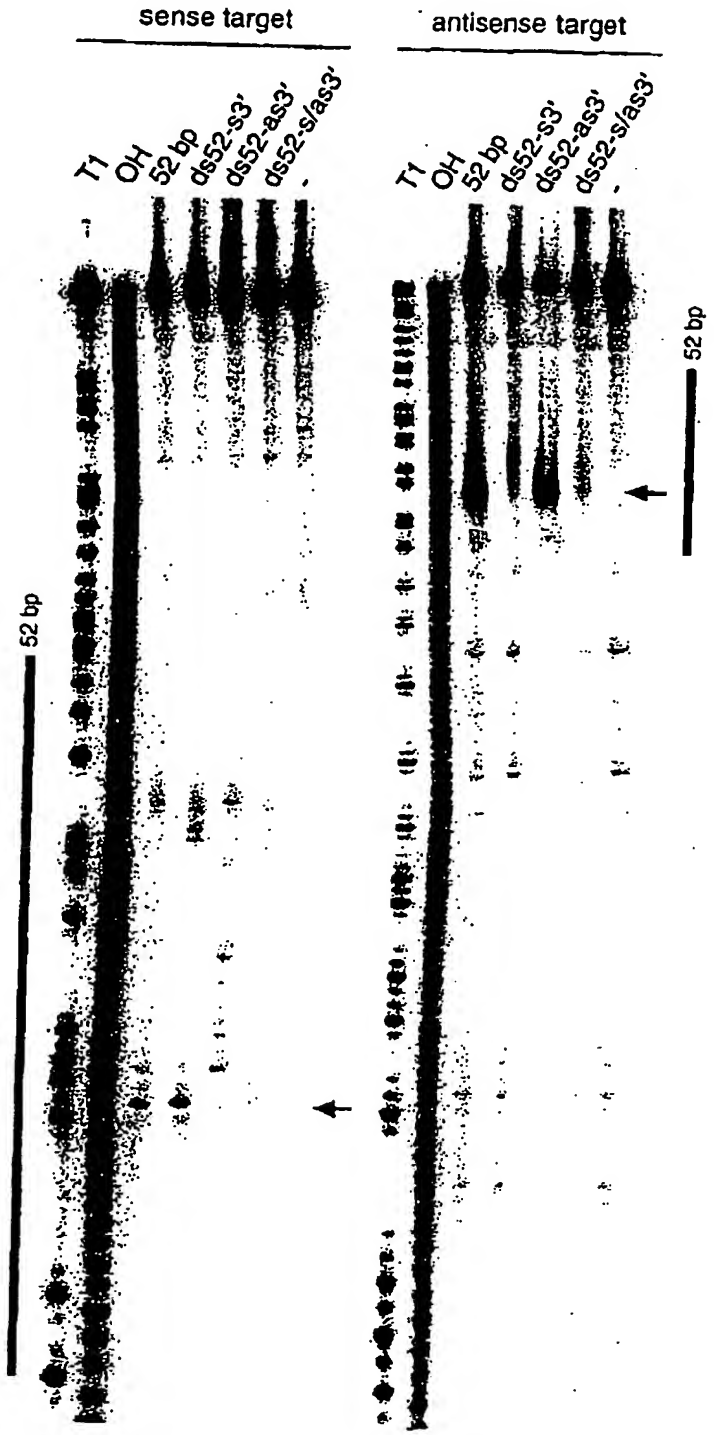


FIGURE 6B


sense target
antisense target
T1
OH
52 bp
ds52-s3'
ds52-as3'
ds52-s/as3'
52 bp

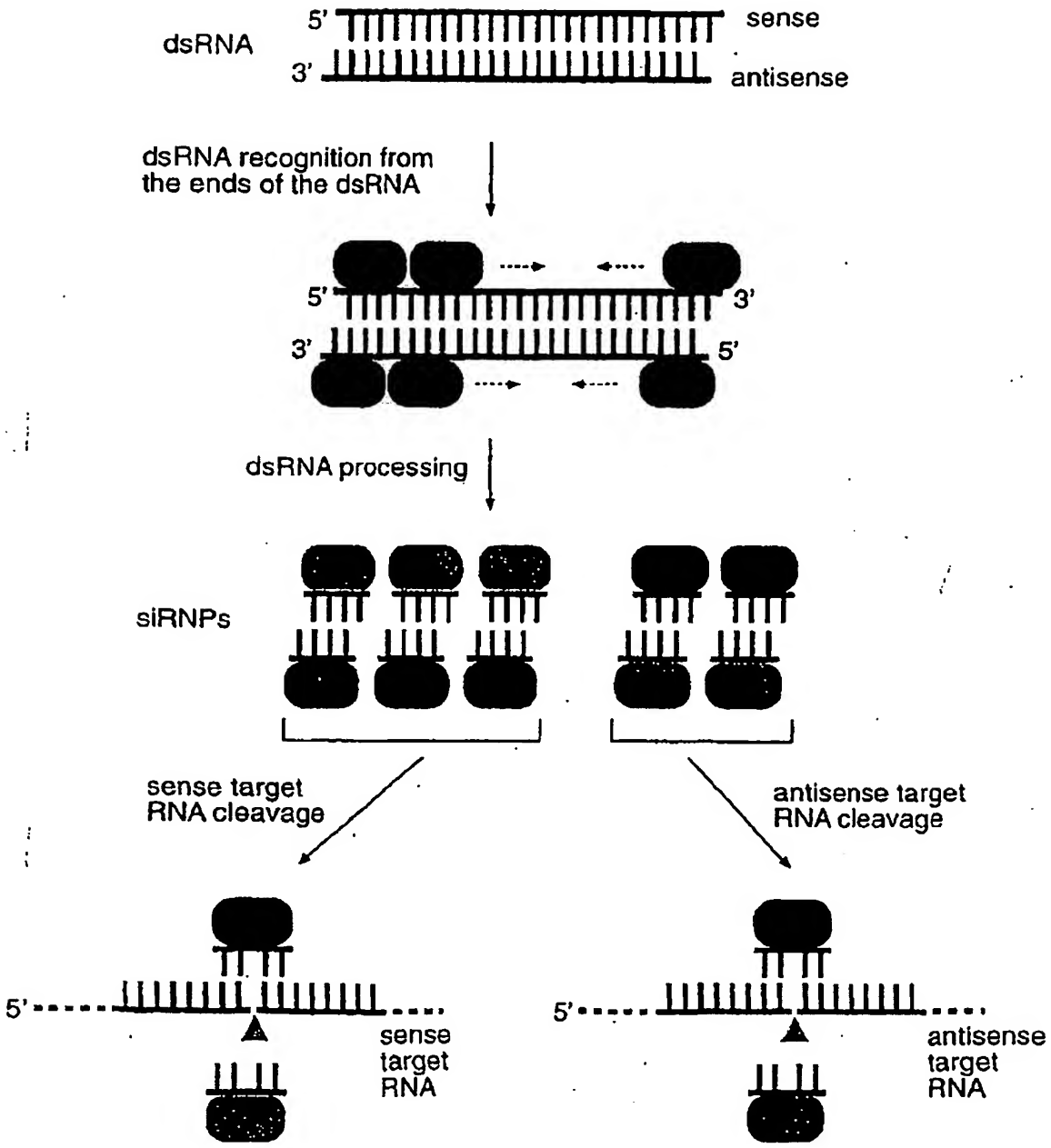
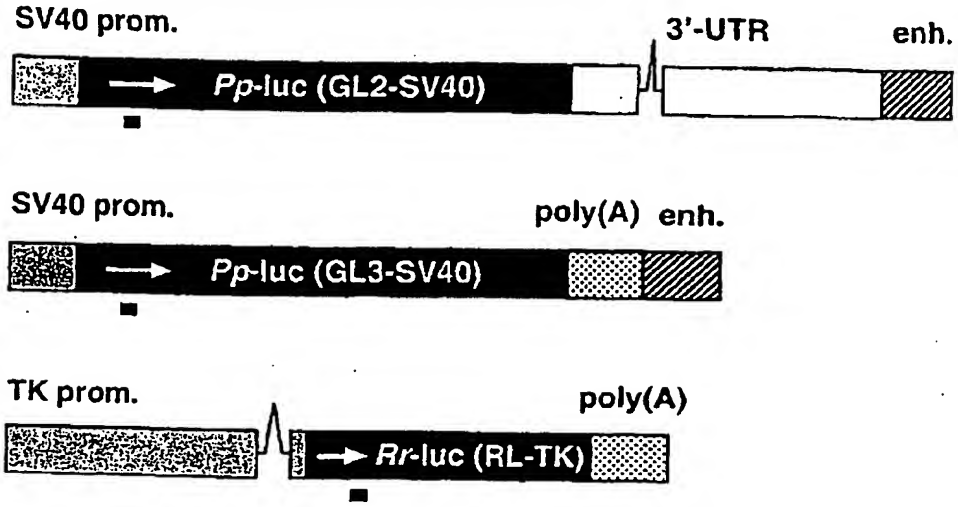

dsRNA
5'
3'
sense
antisense
dsRNA recognition from the ends of the dsRNA
5'
3'
3'
5'
dsRNA processing
siRNPs
sense target RNA cleavage
antisense target RNA cleavage
5'
sense target RNA
5'
antisense target RNA

Figure 7

a



b

siRNA duplex

uGL2	5′ CGUACGCGGAAUACUUCGAUU UUGCAUGCGCCUUAUGAAGCU 5′
GL2	5′ CGUACGCGGAAUACUUCGATT TTGCAUGCGCCUUAUGAAGCU 5′
GL3	5′ CUUACGCUGAGUACUUCGATT TTGAAUGCGACUCAUGAAGCU 5′
invGL2	5′ AGCUUCAUAAGGCGCAUGCTT TTUCGAAGUAUUCCGCGUACG 5′
RL	5′ AAACAUGCAGAAAAUGCUGTT TTUUUGUACGUCUUUUACGAC 5′

Figure 8

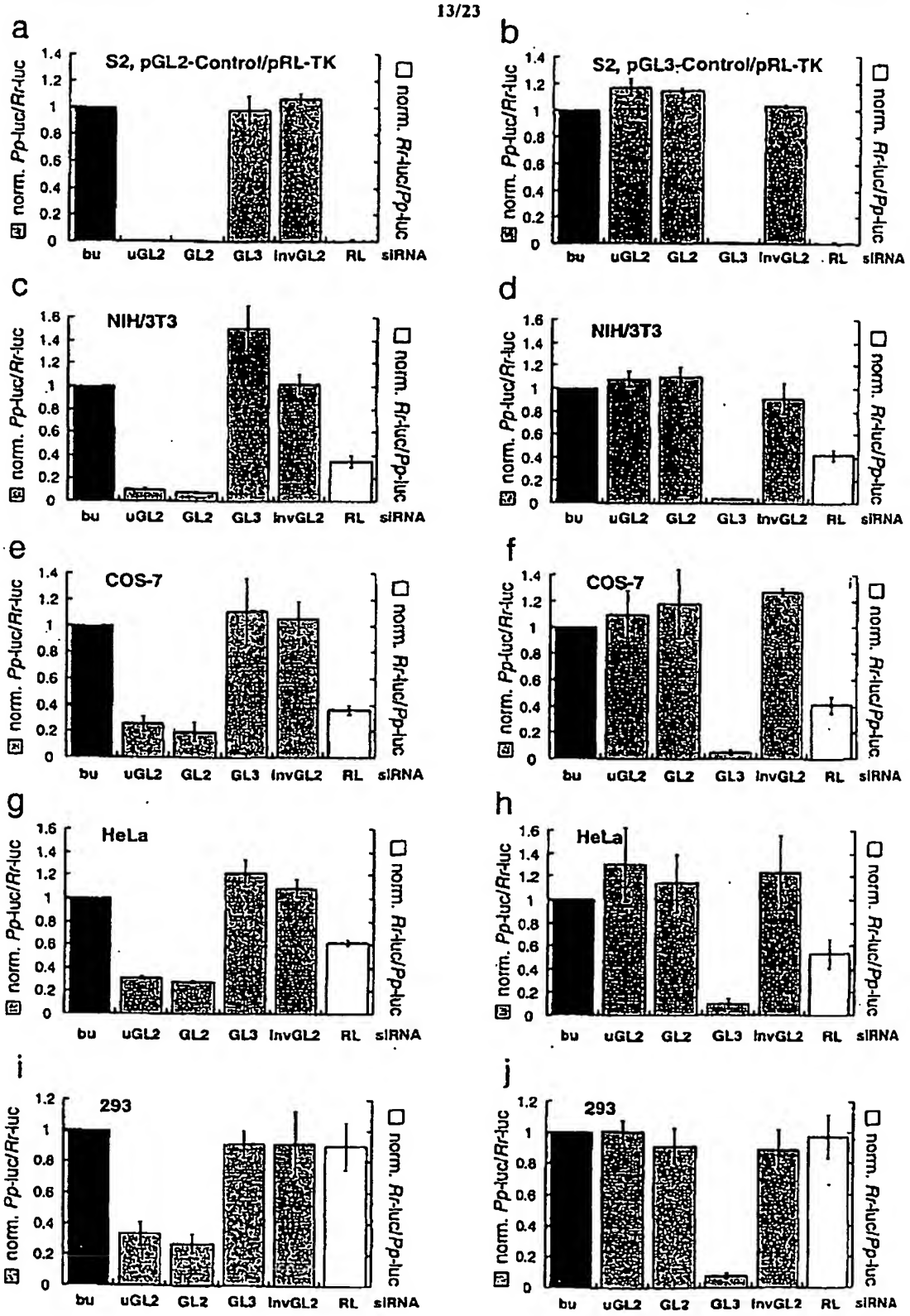

a
S2, pGL2-Control/pRL-TK
b
S2, pGL3-Control/pRL-TK
c
NIH/3T3
d
NIH/3T3
e
COS-7
f
COS-7
g
HeLa
h
HeLa
i
293
j
293
norm. Pp-luc/Rr-luc
norm. Rr-luc/Pp-luc
bu
uGL2
GL2
GL3
invGL2
RL
siRNA

Figure 9

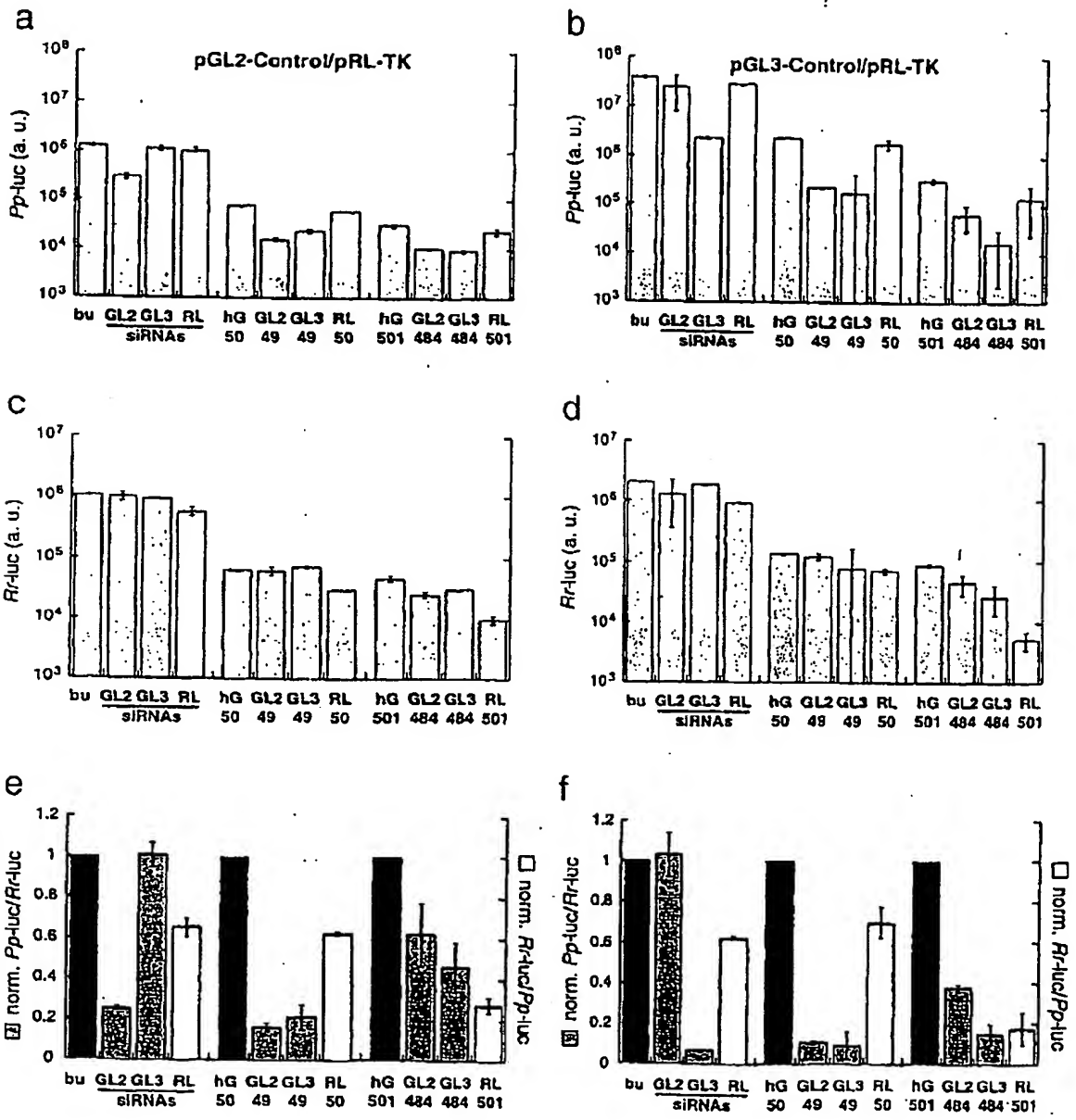
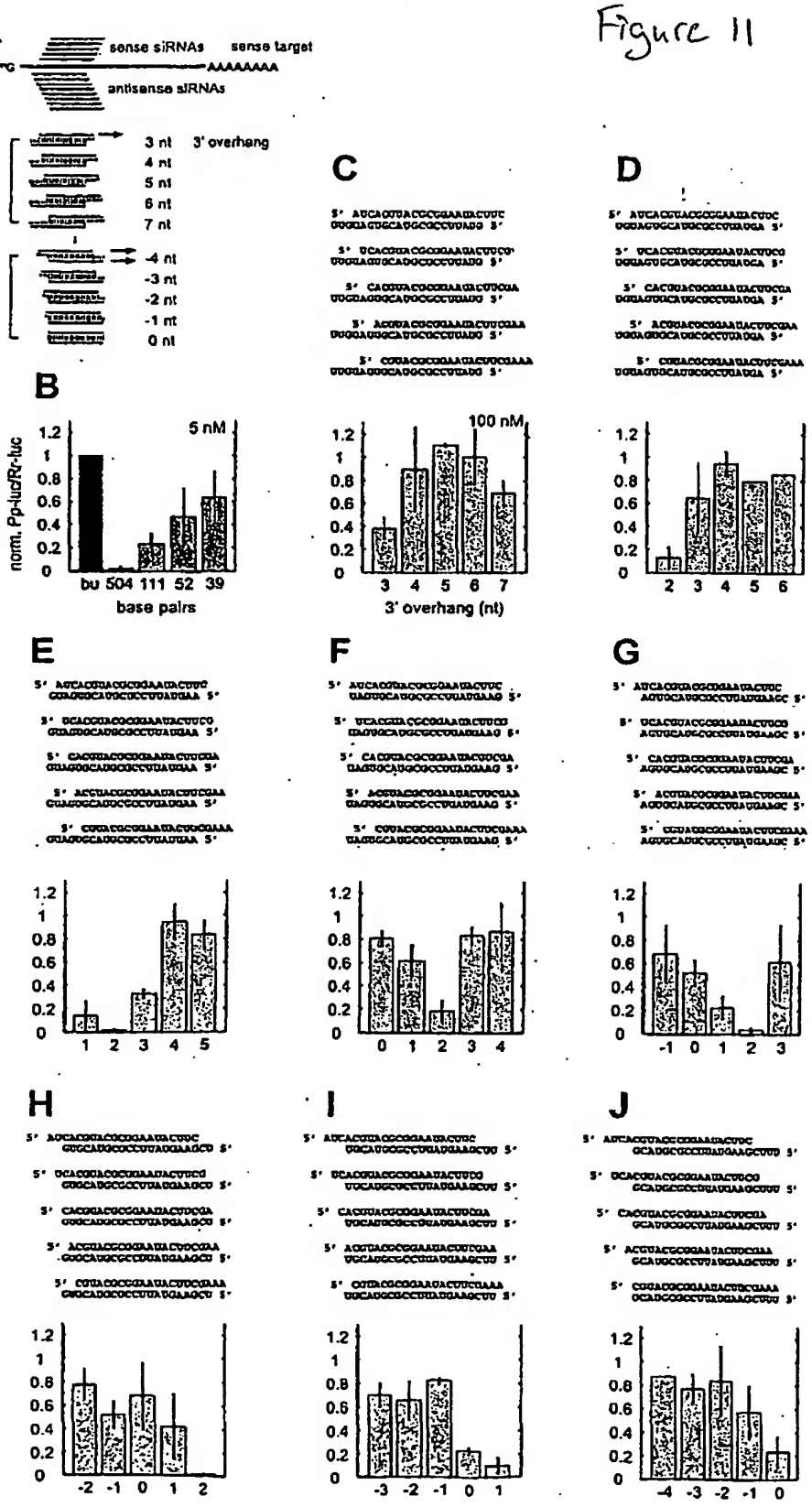

a
pGL2-Control/pRL-TK
Pp-luc (a. u.)
b
pGL3-Control/pRL-TK
Pp-luc (a. u.)
c
Rr-luc (a. u.)
d
Rr-luc (a. u.)
e
norm. Pp-luc/Rr-luc
norm. Rr-luc/Pp-luc
f
norm. Pp-luc/Rr-luc
norm. Rr-luc/Pp-luc
bu GL2 GL3 RL hG GL2 GL3 RL hG GL2 GL3 RL
siRNAs
50 49 49 50 501 484 484 501

Figure 10

Figure 11


A
sense siRNAs
sense target
antisense siRNAs
AAAAAAAA
3 nt
3' overhang
4 nt
5 nt
6 nt
7 nt
-4 nt
-3 nt
-2 nt
-1 nt
0 nt
B
5 nM
norm. Pp-luc/Rr-luc
bu 504 111 52 39
base pairs
C
100 nM
3' overhang (nt)
D
E
F
G
H
I
J

16/23

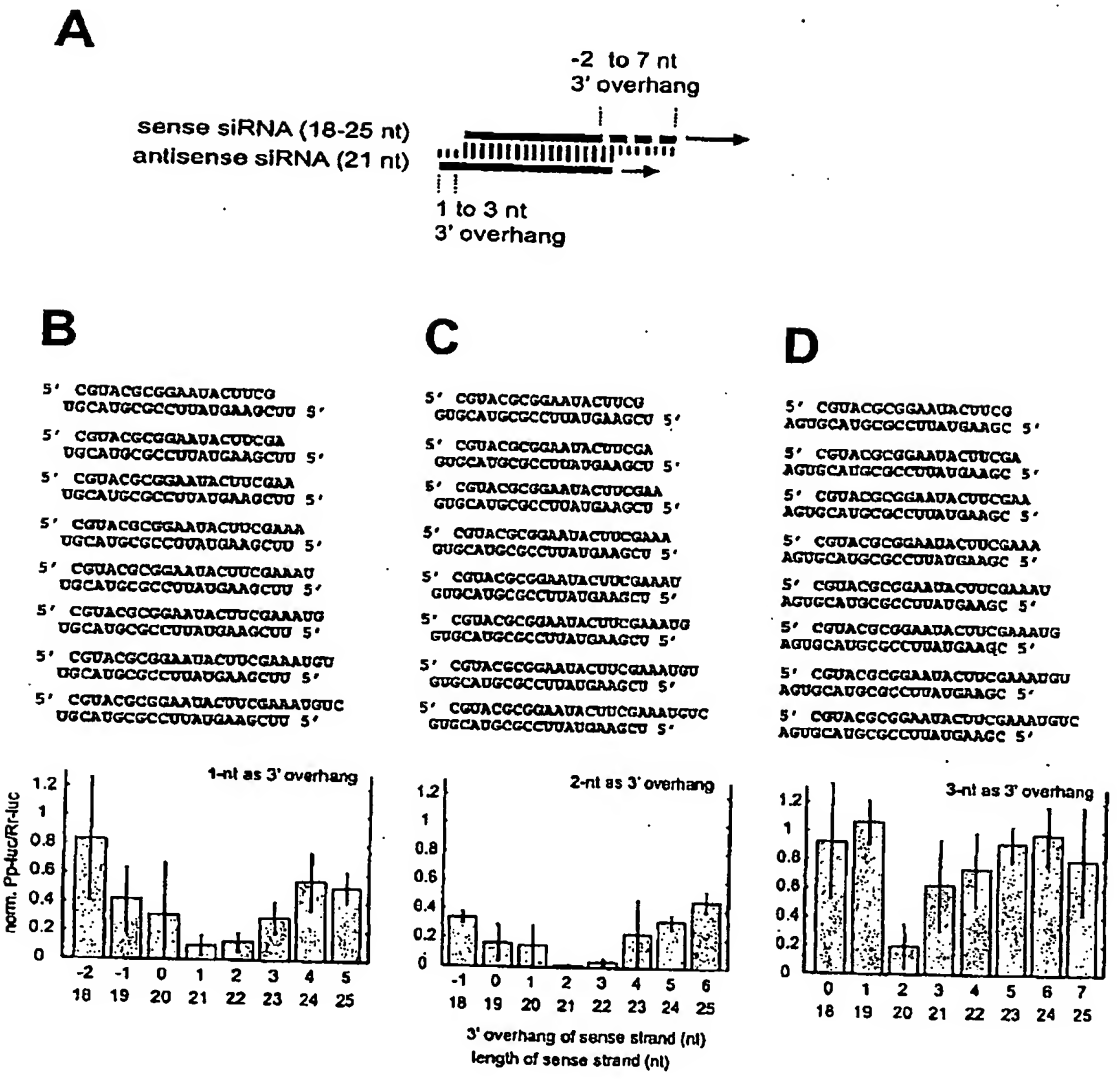

A
-2 to 7 nt
3' overhang
sense siRNA (18-25 nt)
antisense siRNA (21 nt)
1 to 3 nt
3' overhang
B
5' CGUACGCGGAAUACUUCG
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGA
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAA
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAA
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAU
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAUG
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAUGU
UGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
UGCAUGCGCCUUAUGAAGCUU 5'
1-nt as 3' overhang
norm. Pp-luc/Rr-luc
1.2
1
0.8
0.6
0.4
0.2
0
-2 -1 0 1 2 3 4 5
18 19 20 21 22 23 24 25
C
5' CGUACGCGGAAUACUUCG
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAU
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUG
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUGU
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
GUGCAUGCGCCUUAUGAAGCU 5'
2-nt as 3' overhang
1.2
1
0.8
0.6
0.4
0.2
0
-1 0 1 2 3 4 5 6
18 19 20 21 22 23 24 25
3' overhang of sense strand (nt)
length of sense strand (nt)
D
5' CGUACGCGGAAUACUUCG
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGA
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAA
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAA
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAAU
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAAUG
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAAUGU
AGUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
AGUGCAUGCGCCUUAUGAAGC 5'
3-nt as 3' overhang
1.2
1
0.8
0.6
0.4
0.2
0
0 1 2 3 4 5 6 7
18 19 20 21 22 23 24 25

Fig. 13

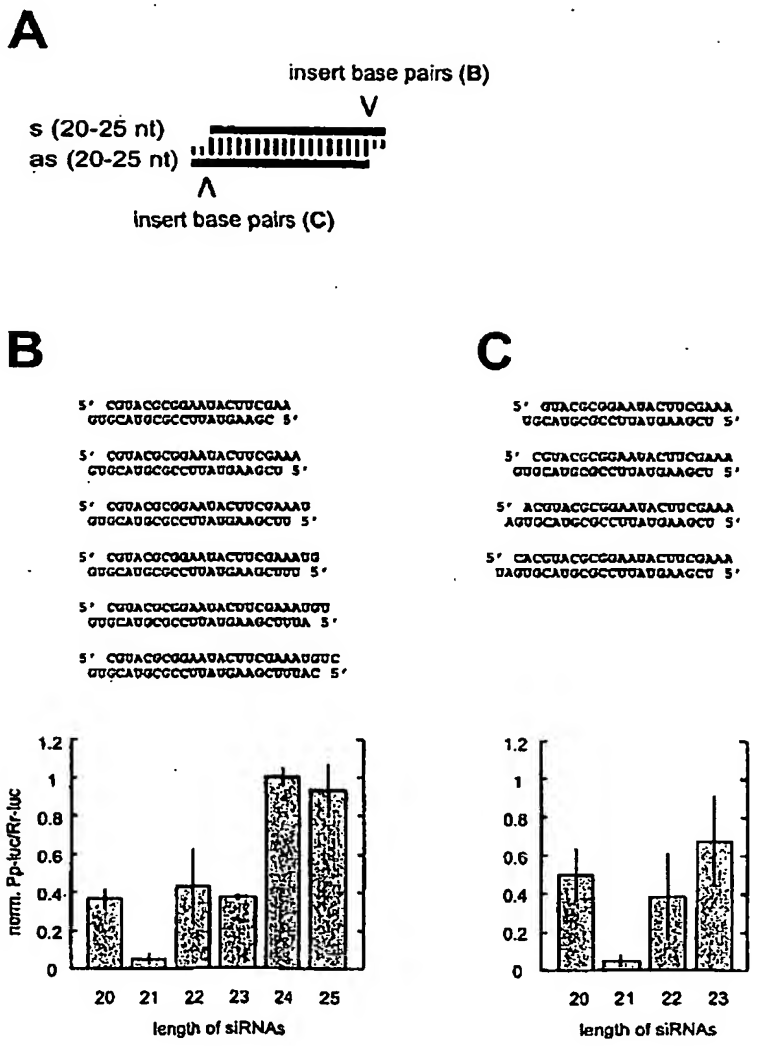

A
insert base pairs (B)
s (20-25 nt)
as (20-25 nt)
insert base pairs (C)
B
5' CGUACGCGGAAUACUUCGAA
GUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAA
CUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAU
GUGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAUG
GUGCAUGCGCCUUAUGAAGCUUU 5'
5' CGUACGCGGAAUACUUCGAAAUGU
GUGCAUGCGCCUUAUGAAGCUUUA 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
GUGCAUGCGCCUUAUGAAGCUUUAC 5'
C
5' GUACGCGGAAUACUUCGAAA
UGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' ACGUACGCGGAAUACUUCGAAA
AGUGCAUGCGCCUUAUGAAGCU 5'
5' CACGUACGCGGAAUACUUCGAAA
UAGUGCAUGCGCCUUAUGAAGCU 5'
norm. Pp-luc/Rr-luc
1.2
1
0.8
0.6
0.4
0.2
0
20 21 22 23 24 25
length of siRNAs
1.2
1
0.8
0.6
0.4
0.2
0
20 21 22 23
length of siRNAs

Fig. 14

s 5' CGUACGCGGAAUACUUCGAAA
as GUGCAUGCGCCUUAUGAAGCU 5'

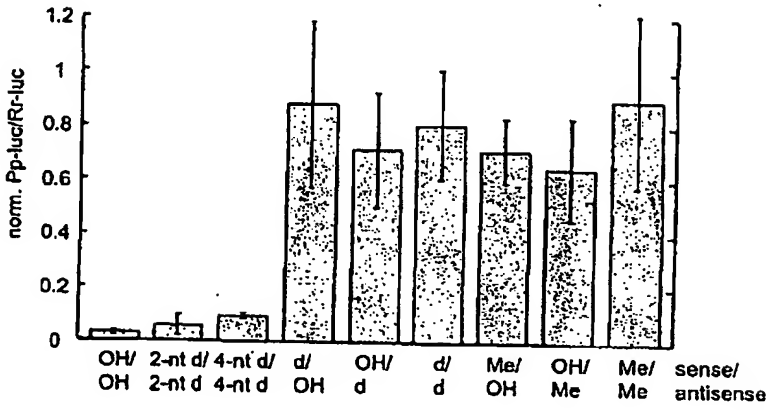
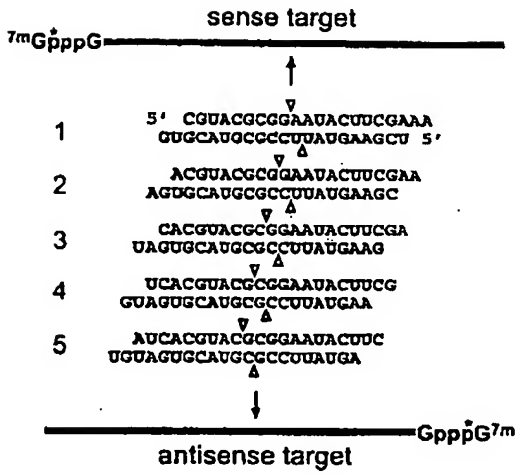


Fig. 15

A


sense target
7mGpppG
1
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
2
ACGUACGCGGAAUACUUCGAA
AGUGCAUGCGCCUUAUGAAGC
3
CACGUACGCGGAAUACUUCGA
UAGUGCAUGCGCCUUAUGAAG
4
UCACGUACGCGGAAUACUUCG
GUAGUGCAUGCGCCUUAUGAA
5
AUCACGUACGCGGAAUACUUC
UGUAGUGCAUGCGCCUUAUGA
GpppG7m
antisense target

B

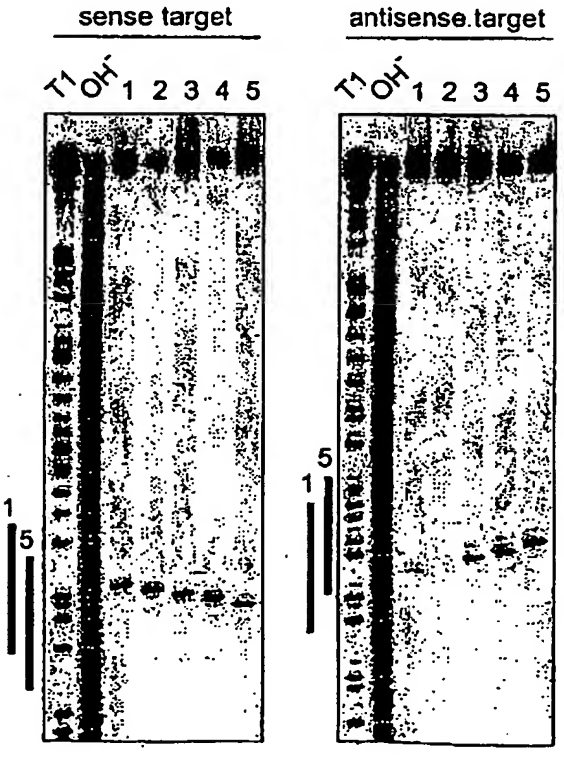

sense target
antisense target
T1
OH-
1 2 3 4 5
1
5

Fig. 16

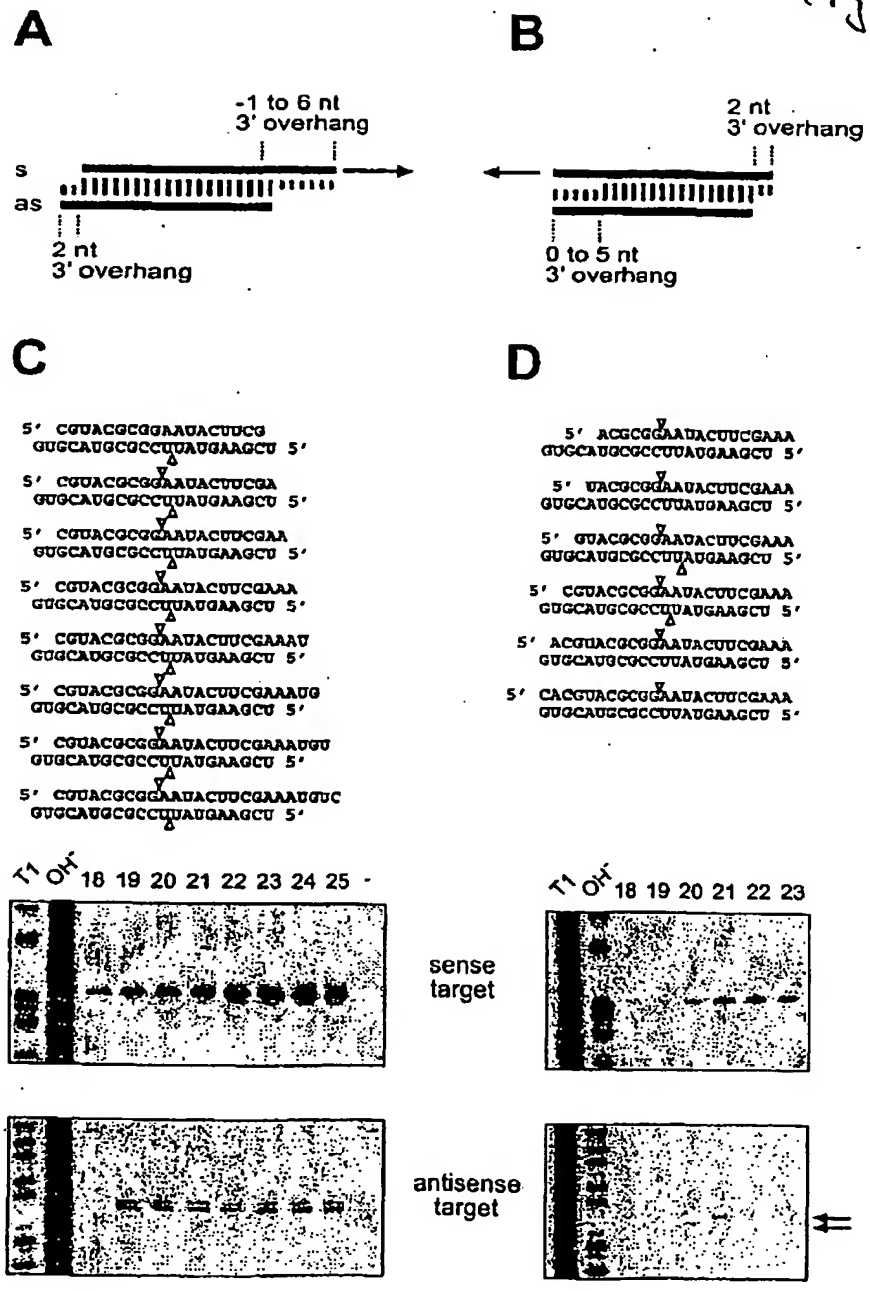

A
-1 to 6 nt
3' overhang
s
as
2 nt
3' overhang
B
2 nt
3' overhang
0 to 5 nt
3' overhang
C
5' CGUACGCGGAAUACUUCG
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAU
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUG
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUGU
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
GUGCAUGCGCCUUAUGAAGCU 5'
D
5' ACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' UACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' GUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' ACGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CACGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
T1 OH- 18 19 20 21 22 23 24 25
T1 OH- 18 19 20 21 22 23
sense target
antisense target

Fig. 17

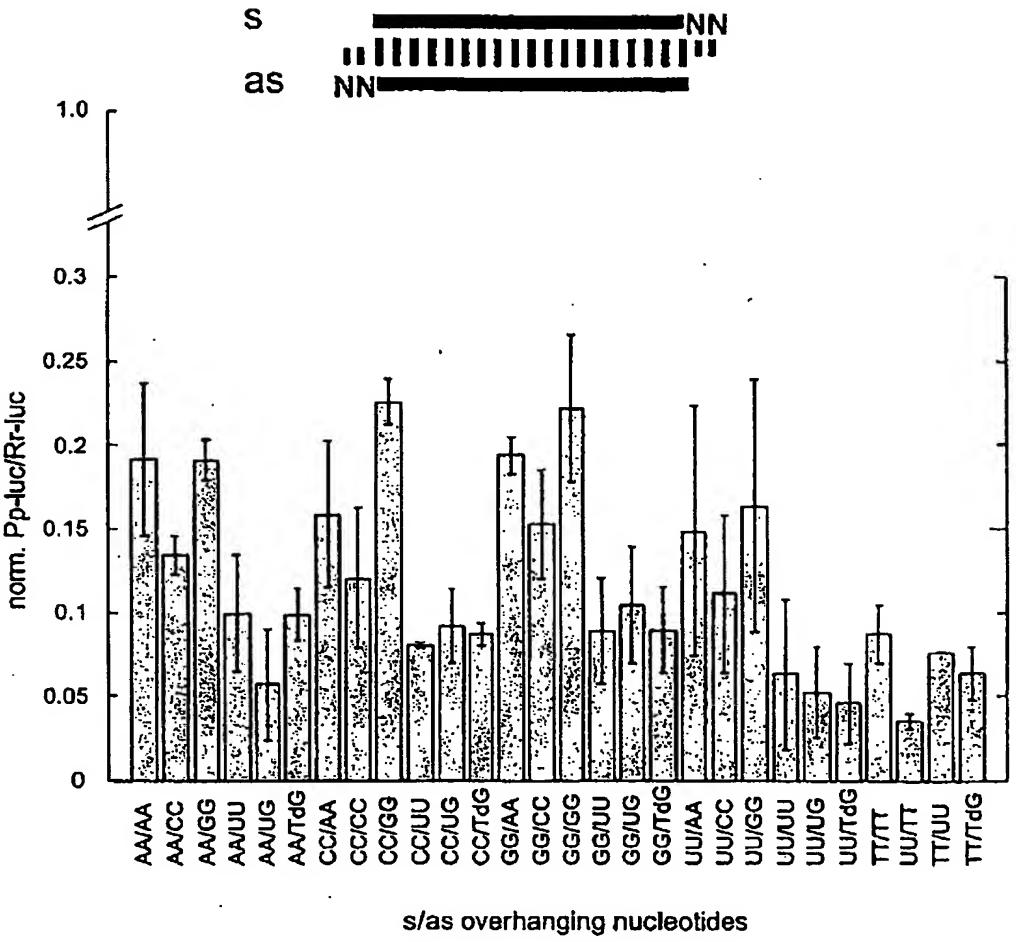

s
NN
as
NN
1.0
0.3
0.25
0.2
0.15
0.1
0.05
0
norm. Pp-luc/Rr-luc
AA/AA
AA/CC
AA/GG
AA/UU
AA/UG
AA/TdG
CC/AA
CC/CC
CC/GG
CC/UU
CC/UG
CC/TdG
GG/AA
GG/CC
GG/GG
GG/UU
GG/UG
GG/TdG
UU/AA
UU/CC
UU/GG
UU/UU
UU/UG
UU/TdG
TT/TT
UU/TT
TT/UU
TT/TdG
s/as overhanging nucleotides

Fig. 18

ref 5' CGUACGCGGAAUACUUCGATT
TTGCAUGCGCCUUAUGAAGCU 5'

1 5' AUGCCGCGGAAUACUUCGATT
TTUACGGCGCCUUAUGAAGCU 5'

2 5' CGUAGCGCGAAUACUUCGATT
TTGCAUCGCGCUUAUGAAGCU 5'

3 5' CGUACGCGAGUUACUUCGATT
TTGCAUGCGCUCAUGAAGCU 5'

4 5' CGUACGCGGAAUUUCCCGATT
TTGCAUGCGCCUUAAAGGGCU 5'

5 5' CGUACGCGGAAUACUUACGTT
TTGCAUGCGCCUUAUGAAUGCG 5'

6 5' CGUACGCGGUAUACUUCGATT
TTGCAUGCGCCAUAUGAAGCU 5'

7 5' CGUACGCGGAUUACUUCGATT
TTGCAUGCGCCUAAUGAAGCU 5'

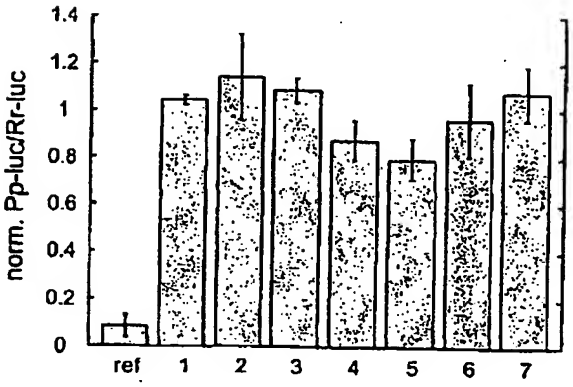
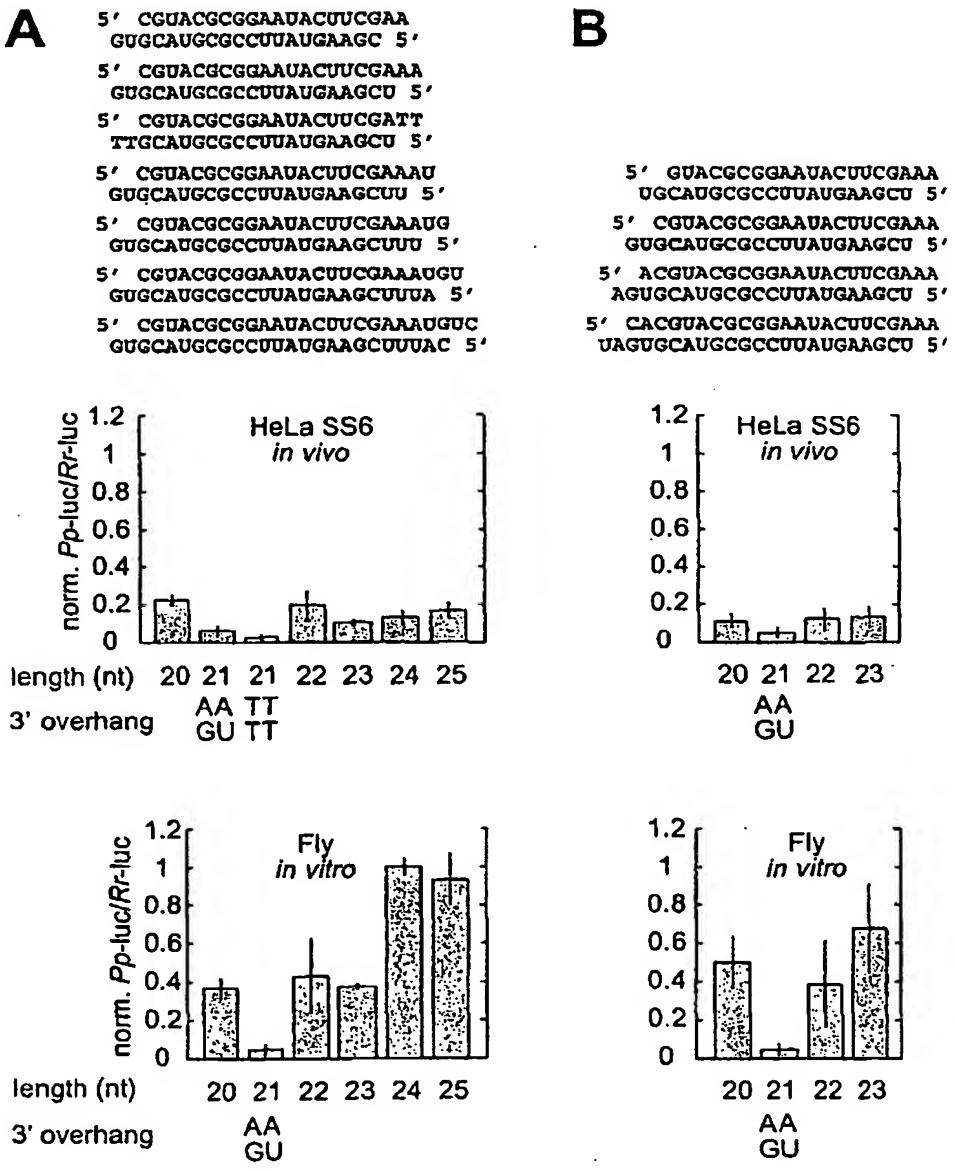


Figure 19


A
5' CGUACGCGGAAUACUUCGAA
GUGCAUGCGCCUUAUGAAGC 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGATT
TTGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAAU
GUGCAUGCGCCUUAUGAAGCUU 5'
5' CGUACGCGGAAUACUUCGAAAUG
GUGCAUGCGCCUUAUGAAGCUUU 5'
5' CGUACGCGGAAUACUUCGAAAUGU
GUGCAUGCGCCUUAUGAAGCUUUA 5'
5' CGUACGCGGAAUACUUCGAAAUGUC
GUGCAUGCGCCUUAUGAAGCUUUAC 5'
B
5' GUACGCGGAAUACUUCGAAA
UGCAUGCGCCUUAUGAAGCU 5'
5' CGUACGCGGAAUACUUCGAAA
GUGCAUGCGCCUUAUGAAGCU 5'
5' ACGUACGCGGAAUACUUCGAAA
AGUGCAUGCGCCUUAUGAAGCU 5'
5' CACGUACGCGGAAUACUUCGAAA
UAGUGCAUGCGCCUUAUGAAGCU 5'
norm. Pp-luc/Rr-luc
HeLa SS6
in vivo
1.2
1
0.8
0.6
0.4
0.2
0
length (nt) 20 21 21 22 23 24 25
3' overhang AA TT GU TT
HeLa SS6
in vivo
20 21 22 23
AA
GU
norm. Pp-luc/Rr-luc
Fly
in vitro
length (nt) 20 21 22 23 24 25
3' overhang AA GU
Fly
in vitro
20 21 22 23
AA
GU

Biochemistry 1996, *35*, 9983−9994

Minor-Groove Recognition of Double-Stranded RNA by the Double-Stranded RNA-Binding Domain from the RNA-Activated Protein Kinase PKR†

Philip C. Bevilacqua and Thomas R. Cech*

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado 80309-0215

Received March 25, 1996; Revised Manuscript Received May 15, 1996®

ABSTRACT: The human double-stranded RNA- (dsRNA) activated protein kinase (PKR) has a dsRNA-binding domain (dsRBD) that contains two tandem copies of the dsRNA-binding motif (dsRBM). The minimal-length polypeptide required to bind dsRNA contains both dsRBMs, as determined by mobility-shift and filter-binding assays. Mobility-shift experiments indicate binding requires a minimum of 16 base pairs of dsRNA, while a minimal-length site for saturation of longer RNAs is 11 base pairs. Bulge defects in the helix disfavor binding, and single-stranded tails do not strongly influence the dsRNA length requirement. These polypeptides do not bind an RNA−DNA hybrid duplex or dsDNA as judged by either mobility-shift or competition experiments, suggesting 2′-OH contacts on both strands of the duplex stabilize binding. Related experiments on chimeric duplexes in which specific sets of 2′-OHs are substituted with 2′-H or 2′-OCH_3 reveal that the 2′-OHs required for binding are located along the entire 11 base-pair site. These results are supported by Fe(II) EDTA footprinting experiments that show protein-dependent protection of the minor groove of dsRNA. The dependence of dsRNA−protein binding on salt concentration suggests that only one ionic contact is made between the protein and dsRNA phosphate backbone and that at physiological salt concentrations 90% of the free energy of binding is nonelectrostatic. Thus, the specificity of PKR for dsRNA over RNA−DNA hybrids and dsDNA is largely due to molecular recognition of a network of 2′-OHs involving both strands of dsRNA and present along the entire 11 base-pair site.

Ribonucleoprotein (RNP) complexes are involved in many biological processes including transcription, posttranscriptional processing, gene regulation, translation, nucleocytoplasmic transport, and mRNA stability. In recent years, the identification of conserved sequences for RNA-binding proteins has led to the description of RNA-binding motifs (RBMs), including the double-stranded RNA- (dsRNA) binding motif (dsRBM) (Mattaj, 1993; Burd & Dreyfuss, 1994). The dsRBM was initially identified as a conserved stretch of 65−68 amino acids on the basis of sequence lignment of functionally diverse proteins from a wide range of organisms (St Johnston et al., 1992). A recent search has identified 44 dsRBM sequences from 27 proteins (Kharrat et al., 1995); these include PKR, the *Drosophila* staufen protein required for mRNA localization in the egg, the *Escherichia coli* dsRNA nuclease RNase III, and the mammalian dsRNA−adenosine deaminases (dsRADs) (Kim et al., 1994; O'Connell et al., 1995; Melcher et al., 1996).

The RNA-binding properties of polypeptides derived from the human dsRNA-dependent protein kinase PKR (also termed dsI or DAI for the dsRNA-activated inhibitor) are studied here. PKR is an interferon-induced, viral-response agent that undergoes dimerization and autophosphorylation in the presence of dsRNA, leading to dsRNA-independent phosphorylation of the eukaryotic translation initiation factor eIF-2 and inhibition of translation. More recent work indicates that PKR is involved in normal control of cell growth and differentiation and in regulation of the transcription of specific genes by dsRNA [reviewed in Clemens (1992), Hovanessian (1993), Mathews (1993), Samuel (1993), and Proud (1995)].

Like other RBMs, the dsRBM is modular and can be found in single or multiple copies in a single protein. PKR contains two tandem, N-terminal copies of the dsRBM, designated dsRBM1 and dsRBM2, and a C-terminal kinase domain (Katze et al., 1991; Feng et al., 1992; Green & Mathews, 1992; McCormack et al., 1992; Patel & Sen, 1992). dsRBM1 closely matches the dsRBM consensus sequence, while dsRBM2 matches the consensus sequence primarily in its C-terminal part (St Johnston et al., 1992). In addition, mutagenesis studies indicate that dsRBM1 is more important than dsRBM2 for dsRNA binding (Green & Mathews, 1992; McCormack et al., 1994; Green et al., 1995; Romano et al., 1995).

Structural details of protein−RNA interaction are well understood for several sequence-specific RBDs. The best characterized complex involves the RNP domain from the spliceosomal protein U1A complexed with a 21-nucleotide RNA hairpin. The crystal structure reveals the RNP making detailed sequence-specific contacts with seven nucleotides in the hairpin loop (Oubridge et al., 1994). Structures of other RNA−protein complexes also reveal sequence-specific interaction with RNA, including a bacteriophage MS2 coat protein−19-nucleotide RNA fragment complex (Valegård et al., 1994), several tRNA synthetase−tRNA complexes (Rould et al., 1989; Ruff et al., 1991), and TAR−arginine and TAR−peptide complexes (Puglisi et al., 1992, 1995; Aboul-ela et al., 1995; Ye et al., 1995). The RNAs in these

† This work is supported by a fellowship to P.C.B. from the Jane Coffin Childs Memorial Fund for Medical Research. T.R.C. is an Investigator of the Howard Hughes Medical Institute and a Professor of the American Cancer Society.

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts*, July 15, 1996.

S0006-2960(96)00725-8 CCC: $12.00

complexes have bulges or loops that can distort the dsRNA helix, opening and widening the usually deep and narrow inaccessible major groove (Weeks & Crothers, 1993). Since the major groove contains most of the sequence-specific information, bulges render the RNA accessible to sequence-specific protein interactions (Mattaj, 1993; Steitz, 1993; McCarthy & Kollmus, 1995).

In contrast to the above examples, undistorted A-form dsRNA has its sequence-rich information buried in the major groove (Saenger, 1984; Steitz, 1993). Indeed, no sequence specificity has been observed in interactions between dsRBDs and RNA *in vitro* (Hunter et al., 1975; Manche et al., 1992; Polson & Bass, 1994; Schweisguth et al., 1994; Bycroft et al., 1995a). Furthermore, PKR does not make important contacts to bases when it binds adenovirus inhibitory RNA (VA RNA$_I$) (Clarke & Mathews, 1995). Recognition of dsRNA is thus likely to be novel and to involve a network of sequence-independent interactions. In this paper, we examine the roles of non-sequence-specific dsRNA functional groups, including 2′-OHs and phosphates, in binding to polypeptides and present a model to account for this binding.

MATERIALS AND METHODS

Expression and Purification of PKR Protein Constructs. C-Terminal deletion protein constructs were prepared without a (His)$_6$ tag, with an N-terminal (His)$_6$ tag, or with a C-terminal (His)$_6$ tag. Protein constructs without a (His)$_6$ tag were a gift (P. DuCharme and S. C. Schultz, personal communication). The cDNA for PKR was obtained from plasmid pB1 Nde P1 KIN (Thomis et al., 1992). Protein constructs with a C-terminal (His)$_6$ tag were prepared as follows. PCR was used to (1) introduce a recognition site for *Eco*RI 5′ to PKR coding sequences and (2) add six histidine codons, alternating between CAC and CAT codons; the stop codon, TAA; and a *Bam*HI site 3′ to PKR coding sequences. Since the coding sequences contain an internal *Eco*RI site, a complete digestion with *Bam*HI was followed by a limited digestion with *Eco*RI to allow for approximately 10% digestion. The PCR fragments were cloned into the T7 expression plasmid PKT7(-H) (S. C. Schultz and T. A. Steitz, personal communication) that had been digested with *Eco*RI and *Bam*HI.

Protein constructs with an N-terminal (His)$_6$ tag were prepared as follows. PCR was used to (1) introduce a recognition site for *Nde*I 5′ to PKR coding sequences and (2) add the stop codon, TAA, and a *Bam*HI site 3′ to PKR coding sequences. The PCR fragment was digested to completion first by *Nde*I and second by *Bam*HI. The fragment was cloned into the T7 expression plasmid pET-14b (Novagen) that contains sequences required for T7 RNA polymerase-driven overexpression, an N-terminal (His)$_6$ tag, and a thrombin restriction site for removal of the (His)$_6$ tag. Most experiments were performed with the N-terminal (His)$_6$ protein constructs. pET-14b offers the advantage that the cloning sites, *Bam*HI and *Nde*I, do not occur in the PKR coding region, allowing for rapid cloning. Sequences were confirmed by dideoxy sequencing.

Optimal expression of C-terminal deletion protein constructs was in *E. coli* strain BL21(DE3). Cells were grown at 37 °C for 12 h in LB media supplemented with 20 mM potassium phosphate (pH 7.8), 5 mM glucose, and 200 μg/mL ampicillin. Cells (5 mL) were centrifuged at 6000 rpm for 10 min, resuspended in 5 mL of LB, and diluted into 750 mL of the above media without glucose. Growth was continued at 37 °C with vigorous shaking in 2-L baffle flasks until OD_{600} = 0.3. The flasks were then shaken at 22 °C until OD_{600} = 0.6−0.8. Expression was induced by the addition of IPTG to 0.4 mM, and growth continued an additional 8 h at 22 °C. Cells were pelleted by centrifugation (10 min at 4000 rpm in a Beckman JA-10 rotor at 4 °C) and stored overnight at −20 °C. All subsequent purification steps were performed at 4 °C.

Protein was soluble and purified by native methods. Cells were resuspended in 20 mL of ice-cold sonication buffer [SB: 50 mM Hepes (pH 7.0), 700 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, 0.1 mM PMSF, and 0.05 mM benzamidine]. Lysozyme was added to 5 mg/mL, and the cells were incubated for 30 min with rotation, followed by sonication. The lysate was cleared by addition of one-tenth volume of 5% polyethyleneimine (pH 9.0; 25 000−50 000 average MW, Aldrich), inverted, incubated on ice for 15 min, and centrifuged (15 min at 10 000 rpm in a Beckman JA-20 rotor) (Schmedt et al., 1995). The supernatant was centrifuged (30 min at 38 000 rpm in a Beckman 70Ti rotor), collected, and subsequently rotated for 30 min with 4 mL of a 50% slurry of Ni^{2+}−nitrilotriacetic acid−agarose resin (Qiagen) previously equilibrated in SB. Imidazole (pH 7.0) was added to 1 mM, and the slurry was incubated another 30 min with rotating. The resin was then pelleted by centrifugation in a table-top swinging bucket rotor for 5 min, and the supernatant was removed. The resin was washed three times by resuspending in 40 mL of ice-cold SB plus 1 mM imidazole, rotating for 15 min, and pelleting. Washing was done an additional four times with wash buffer [WB: 50 mM Hepes (pH 7.0), 700 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, and 30 mM imidazole (pH 7.0)]. Protein was eluted by resuspending the resin in 3 mL of elution buffer [EB: 50 mM Hepes (pH 7.0), 700 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, and 300 mM imidazole (pH 7.0)], rotating for 15 min, pelleting, and combining the supernatants a total of four times. The supernatant was concentrated to 2 mL by ultrafiltration in a Centriprep-10 (10 kDa cutoff) (Amicon) and exchanged three times in storage buffer [StB: 25 mM Hepes (pH 7.0), 50 mM NaCl, 5% glycerol, 2.0 mM DTT, and 0.25 mM EDTA] by resuspending in 15 mL of StB and concentrating to 2 mL each time. Protein was stored at 4 °C. Glycerol was removed prior to Fe(II) EDTA mapping experiments by exchanging the buffer into StB minus glycerol.

The purity of recombinant C-terminal truncated PKR was estimated to be >90% from overloaded Coomasie blue stained protein gels. The concentration of protein was generally determined by the relative Coomasie blue staining on protein gels with lysozyme standards, while the concentration of p24, used to obtain the data in Table 1 and in Figure 1, was determined spectrophotometrically (Gill & von Hippel, 1989). In control experiments, the N-terminal (His)$_6$ tag in 184 and 220 amino acid proteins was removed by a thrombin digest as per manufacturer's instructions (Novagen).

Preparation of RNAs, DNAs, and Chimeras. TAR and dsTAR were prepared by T7 transcription reactions (5 mL) containing 40 mM Tris (pH 7.6), 15 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 1 mM each nucleoside triphosphate, 0.75 μM annealed promoter-template, and 5000 units/mL phage T7 RNA polymerase (Milligan & Uhlenbeck, 1989) and incubated at 37 °C for 2 h. The promoter sequence was

the 23mer: 5'GAAATTAATACGACTCACTATAG3'. Samples were purified in 6% acrylamide gels/8 M urea, visualized by UV shadowing, excised from the gel, and eluted by crushing the gel slice and soaking overnight at 4 °C in TEN_{250} [TEN_{250}: 10 mM Tris (pH 7.5), 1 mM EDTA, and 250 mM NaCl]. RNA was concentrated by ethanol precipitation, washed with 70% ethanol, and quantitated spectrophotometrically.

All other RNA, DNA, and chimeric oligomers were prepared by solid-phase synthesis and deblocked as previously reported (Zaug et al., 1994). Oligomer sequences are found in the appropriate figure or figure caption. Positions of 2'-deoxy- or 2'-methoxy-substituted sugars were confirmed by a limited hydrolysis of the 5'-^{32}P-labeled chimera, followed by running a sequencing gel.

5'-^{32}P-labeled RNAs were generated by treatment with calf intestinal phosphatase (for T7 transcripts only), reacted with polynucleotide kinase and [γ-^{32}P]ATP, repurified by gel electrophoresis, excised from the gel, eluted overnight in TEN_{250} at 4 °C, ethanol precipitated, and resuspended in TE [10 mM Tris (pH 7.5) and 0.1 mM EDTA]. Labeled duplexes were prepared by annealing the 10 nM 5'-^{32}P-labeled strand with a 20-fold excess of complementary strand in TEN_{100} [TEN_{100}: 10 mM Tris (pH 7.5), 1 mM EDTA, and 100 mM NaCl] at 95 °C for 3 min and cooling on the bench for 10 min. Annealed duplexes were stored at −20 °C and used immediately after thawing at 22 °C. Control experiments showed no binding of protein to ssRNA.

Binding Assays. Dissociation constants were determined by either native-gel mobility-shift assays or by filter binding. Duplex RNA was 5'-^{32}P-labeled and present in limiting concentration relative to protein concentrations. Samples were prepared in standard binding buffer [BB: 25 mM Hepes (pH 7.5), 10 mM NaCl, 5% glycerol, 5 mM DTT, 0.1 mM EDTA, and 0.1 mg/mL herring sperm DNA (Sigma)]. Herring sperm DNA was fragmented by sonication to an average length of 3−4 kb, boiled for 10 min, and placed immediately on ice. Herring sperm DNA, or tRNA, as appropriate, was used in each mobility-shift assay to prevent sticking of the complex in the wells of the gel. The two binding methods gave similar results; however, the mobility-shift assay offered the advantage that multiple-protein−RNA complexes, important to the interpretation of the data presented here, could be directly visualized. In addition, filter binding experiments with short substrates suffered from poor retention efficiency, especially at high salt concentrations.

For the mobility shift assay, binding reactions were loaded onto a running 10% (79:1 acrylamide/bis) native gel. The gel and the running buffer contained 0.5× TBE [50 mM Tris base, 41.5 mM boric acid, and 1 mM EDTA, (final pH 8.3)]. Electrophoresis was performed for 1.5 h at 19 V/cm, at 22 °C.

Filter binding experiments were performed in a 96-well dot blot apparatus essentially as described (Wong & Lohman, 1993; Weeks & Cech, 1995) with the following differences. Nitrocellulose (Schleicher & Schuell) and Hybond N+ membranes were equilibrated in BB for 30 min at 22 °C. Wells were washed with 100 μL of BB, after which four reactions (10 μL each) were filtered. Wells were immediately washed with 100 μL of ice-cold BB.

Dissociation constants for chimera and salt dependence experiments were determined by quantifying the fraction (θ) of RNA bound with a PhosphorImager (Molecular Dynamics) and fitting by nonlinear least squares as a function of total PKR concentration (eq 1), where ϵ is the observed

$$\theta = \epsilon \frac{[PKR]}{[PKR] + K_d} \quad (1)$$

maximum fraction bound (typically ≈0.8) and K_d is the dissociation constant. Control experiments were performed with 5 and 30 min of incubation of the binding reaction prior to loading the gel and gave similar results with the optimal fraction bound occurring at 5 min. All mobility-shift assays were thus performed with 5 min of incubation prior to loading the gel. For unsubstituted and MID-substituted chimeric duplexes which gave two band shifts, K_ds were calculated by treating bound RNA as a single species equal to the sum of both bands.

Dissociation constants for binding to TAR and dsTAR were determined by using a two-site binding model, quantifying the fraction of RNA bound in complex 1 (θ_1) and complex 2 (θ_2) with a PhosphorImager, and simultaneously fitting (θ_1) and (θ_2) to eqs 2 and 3, The interaction free energy between the two sites, a measure of cooperativity, was determined by eq 4, where the last term arises from statistical features due to a reduced number of sites for binding of the second protein (Cantor & Schimmel, 1980).

$$\theta_1 = \frac{[PKR]K_{d2}}{[PKR]^2 + [PKR]K_{d2} + K_{d1}K_{d2}} \quad (2)$$

$$\theta_2 = \frac{[PKR]^2}{[PKR]^2 + [PKR]K_{d2} + K_{d1}K_{d2}} \quad (3)$$

$$\Delta G_i = +RT \ln \frac{K_{d2}}{K_{d1}} - RT \ln 4 \quad (4)$$

Fe(II) EDTA Mapping. Labeled chimeric duplexes were prepared by annealing a 5'-^{32}P-labeled strand with excess complementary strand, as described above. The top strand has a single-stranded tail 5' to a 22 base-pair core, with the tail serving as an internal control. Oligomer sequences are found in the caption to Figure 6. Mapping conditions were adapted from published methods (Tullius & Dombroski, 1986). Protein without any glycerol was added and incubated for 5 min at 22 °C and 5 min on ice. $(NH_4)_2$-$Fe^{II}(SO_4)_2 \cdot 6H_2O$−$Na_2EDTA$, sodium ascorbate, and H_2O_2 were freshly prepared and added sequentially (1 μL each; 10 μL total volume) at final concentrations of 2 mM/4 mM, 10 mM, and 0.1%, respectively, and incubated on ice for 1 min. [In the absence of protein, similar amounts of RNA cleavage (≈20%) occurred at 1, 2, 10, and 30 min at 22 °C, suggesting 1 min is sufficient to obtain maximal cleavage.] Thiourea (10 mM) was added to quench the reaction. Five microliters of a formamide/0.1% SDS loading buffer was added that included 4 μM labeled strand, now unlabeled. Unlabeled strand was added to dissociate the 5'-^{32}P-labeled strand from the duplex since the duplex is of sufficient stability to remain partially formed on the denaturing gel. The mixture was heated to 65 °C for 3 min and put on ice. A 3 μL portion of the quenched reaction was loaded on a 25% (20:1 acrylamide/bis) gel/6 M urea/1× TBE that had been preelectrophoresed for a minimum of 2 h at 75 W.

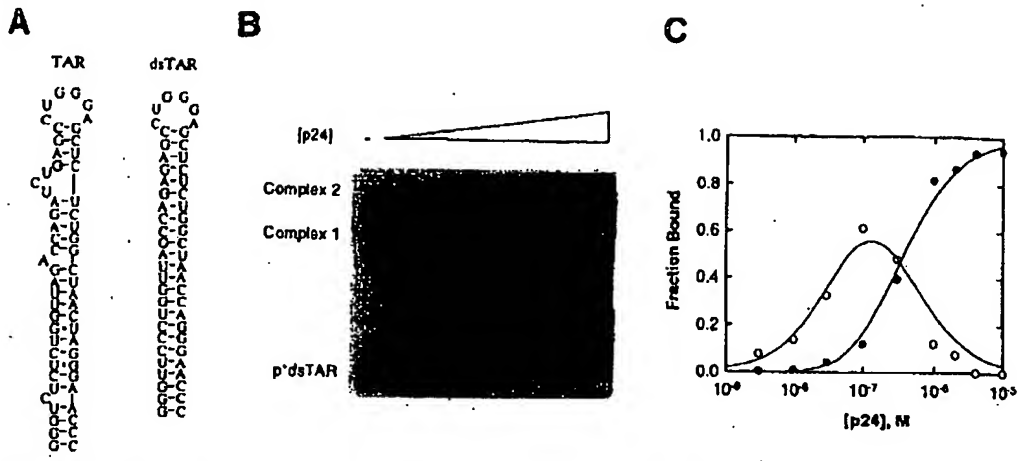


FIGURE 1: Native gel mobility shift for p24 binding to dsTAR. (A) Secondary structures for TAR and dsTAR (Celander & Cech, 1990). (B) Native-gel mobility-shift experiment for p24 binding to trace amounts of 5′-^{32}P-labeled dsTAR RNA. Experiments were in the presence of 0.1 mg/mL ssDNA. Concentrations of p24 used were 0, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 2, and 4 μM. Protein binding to dsTAR resulted in two complexes: Conditions were as described in the text except that samples were loaded after 1 h of preincubation at room temperature onto a 5% (79:1 acrylamide/bis) native gel. (C) Plot of fraction of RNA bound in complex 1 (○) and complex 2 (●) for p24 binding to dsTAR. Fits are to eqs 2−4 and give values of K_{d1} = 0.05 μM, K_{d2} = 0.3 μM, and ΔG_i° = +0.3 kcal/mol (Table 1).

Marker lanes were run in Fe(II) EDTA mapping experiments. G sequencing lanes were prepared by limited hydrolysis with RNase T_1 (and without RNase T_1 as a control), and all-nucleotide sequencing lanes were prepared by treatment with alkali (Donis-Keller et al., 1977).

Computer-Generated Models. A-form RNA coordinates were generated using Insight II molecular modeling software (Biosym Technologies).

RESULTS

Effect of the (His)$_6$ Fusion Tag on Binding. To determine whether use of the (His)$_6$ tag affected the outcome of these experiments, the tag was removed by a thrombin digest. (His)$_6$-free proteins showed identical K_ds, RNA length requirement, and RNA−DNA hybrid band shifts as N-terminal (His)$_6$ tag proteins. The (His)$_6$ tag was not removed for most experiments presented.

A Model System To Study RNA−dsRBD Interactions: Minimum-Length Polypeptides and a Binding Assay. RNA substrates with and without bulges were prepared. dsTAR is a double-stranded version of TAR with a 24 base-pair stem in which the three bulges are deleted and G·U wobble pairs converted to G·C base pairs (Figure 1A). We chose TAR and dsTAR as model RNAs since TAR has been reported to both activate and inactivate PKR depending on TAR concentration, suggesting TAR can bind to PKR (Gunnery et al., 1990, 1992; Roy et al., 1991; Maitra et al., 1994). Also, the TAR RNA-binding protein (TRBP), which has three dsRBMs (Kharrat et al., 1995), binds tightly to TAR RNA and dsRNAs (Gatignol et al., 1991, 1993; Park et al., 1994). These RNAs are able to support dsRNA-specific binding (Figure 1B,C; Table 1).

In order to find a minimal-length polypeptide to study, a number of C-terminal truncated constructs were examined for binding (Figure 2). Constructs that were truncated at or before residue 100 did not give binding that was specific to dsTAR over all-DNA versions of TAR (dTAR). The minimal polypeptide examined that gave RNA-specific binding was 110 amino acids in length; its binding to dsTAR, however, was very weak (Figure 2). The minimal polypeptide that gave strong RNA-specific binding as assayed by either native-gel or filter-binding experiments was 184 residues in length and contained both dsRBM1 and dsRBM2 (Figure 2). These observations are consistent with a report that a construct with residues 1−129 gave no detectable dsRNA binding but a construct with residues 1−170 bound dsRNA (Patel & Sen, 1992). The polypeptides discussed in the remainder of this paper, p20 and p24 as well as their (His)$_6$-tagged analogs, are 184 and 220 residues in length. These polypeptides contain the same PKR amino acids as previously reported constructs (Green & Mathews, 1992; Manche et al., 1992). A 1−243 truncated construct bound RNA with similar affinity as full-length PKR with a catalytic point mutation (McCormack & Samuel, 1995), suggesting C-terminal truncated constructs retain wild-type RNA-binding activity. A longer polypeptide of 280 residues, extending to the kinase domain, bound 22-base pair dsRNA but gave complex mobility shifts with multiplets of four or more bands and was not further investigated (Figure 2). Stable RNA binding by the dsRBD from PKR requires both dsRBM1 and dsRBM2.

Table 1: Effects of Bulges and Competitor on RNA Binding to p24[a]

RNA	competitor (0.1 mg/mL)	K_{d1} (μM)	K_{d2} (μM)	ΔG_i° (kcal/mol)
dsTAR	ssDNA	0.05	0.3	+0.3
dsTAR	tRNA^{Phe}	0.4	0.3	−1
TAR	ssDNA	3	0.07	−3.1
TAR	tRNA^{Phe}	6	1	−1.7

[a] Data are fit to a two-step random-order binding mechanism (see Materials and Methods). According to this model, K_{d1} reflects binding of one protein to RNA and K_{d2} reflects binding of a second protein to RNA. ΔG_i° is an interaction free energy and estimates the cooperativity of protein binding to RNA, where negative values indicate positive cooperativity. Uncertainties are estimated at 30% in K_ds and 5% in ΔG_i°s. There was no detectable binding to an all-deoxy version of TAR, dTAR, under identical conditions.

Effects of RNA Structure and Length on dsRBD Binding. Initial experiments compared binding of p24 to limiting amounts of 5′-^{32}P-labeled RNA in the presence of single-stranded DNA (ssDNA) and tRNA competitors. Binding of p24 to dsTAR or TAR gave rise to two shifted bands of different mobility (e.g., Figure 1B). The fast-mobility band,

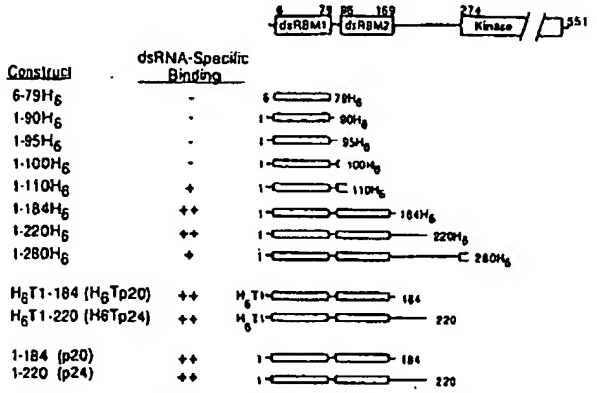


FIGURE 2: Polypeptide construct schematic. (Top) Location of dsRBM1, dsRBM2, and the kinase domain in the 551 residue PKR. dsRBM1 contains amino acids 6–79 and dsRBM2 contains amino acids 96–169 (St Johnston et al., 1992). The kinase domain resides in the C-terminal portion of PKR and contains the 11 submotifs conserved among protein kinases, with domain I starting at residue 274 (Hanks et al., 1988; McCormack et al., 1992). (Bottom) Protein constructs examined for dsRNA binding in this study. Shown are the N- and C-terminal residues of the construct, the presence of any hexahistidine tags (H_6), the presence of any thrombin cleavage sites (T), and appropriate abbreviations. dsRNA-specific binding means binding specific to TAR and dsTAR RNA over an all-DNA version of TAR (dTAR) and is indicated by (−) for no detectable specific binding and (+) for weak ($K_d \gg 1\ \mu M$) and (++) for strong ($K_d < 1\ \mu M$) binding.

complex 1, is an intermediate doublet that formed at low concentrations of p24 and was converted to the slow-migrating complex 2 at high concentration of p24. The doublet nature of complex 1 suggests a minimum of two distinct binding sites; thus a two-site random-order model was chosen to fit this data, in which one protein binds to RNA to give complex 1, followed by binding of a second protein to give complex 2 (Materials and Methods). According to this random-order binding mechanism, formation of complexes 1 and 2 is described by dissociation constants K_{d1} and K_{d2} and an interaction free energy, ΔG°_{i}, that describes any cooperativity for binding of the second protein (Table 1). Complex 2 was resistant to the nonspecific protein competitor bovine serum albumin (BSA 0.5 mg/mL), suggesting complex 2 is not simply due to protein−protein aggregation (data not shown).

Table 1 summarizes the effects of adding bulges in the RNA substrate (i.e., TAR RNA) and varying the competitor. Two trends may be observed: (1) formation of complex 1 is disfavored by bulges and tRNA competitor, and (2) the interaction free energy is largest for the weakest binding combinations. The first trend is consistent with the protein−dsRNA interactions in complex 1 being weakened by bulges and subject to competition by tRNA. [In related experiments, tRNA was found to compete weakly for p20 binding to 85 base-pair dsRNA (Schmedt et al., 1995).] In addition, a stronger interaction free energy for proteins in the presence of bulges and tRNA competitor suggests that complex 2 is not as strongly affected by these factors as complex 1. The second trend is consistent with a second p24 protein binding in a cooperative fashion. This cooperativity could arise from favorable protein−protein interactions on the dsRNA, from an RNA conformational change induced by binding of the first protein, or both. One plausible RNA conformational change would involve the TAR RNA adopting a more uniform double-stranded conformation upon binding of the

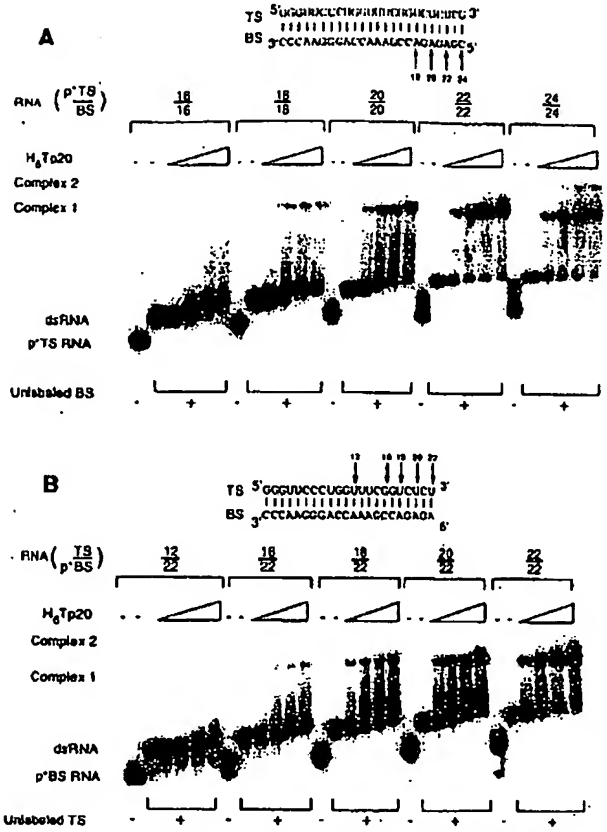


FIGURE 3: RNA length and single-stranded tail dependence. Binding of H_6Tp20 to dsRNAs of discrete length. (A) Mobility-shift experiment of H_6Tp20 binding to trace amounts of dsRNA of varying length (bp). The top-strand (TS) oligomer was 5′-^{32}P-labeled and annealed to excess amounts of unlabeled bottom-strand (BS) oligomer. Formation of dsRNA was confirmed by a microshift of p*TS upon addition of BS. [Compare (−) and (+), the first and second lanes of each RNA length set, respectively.] Sequences of TAR-derived duplexes of 18–24 base pairs are shown at the top. The sequence of 16 bp is (TS) 5′GGGUUCCCUGGUUAGC3′ and (BS) 5′GCUAACCAGGGAACCC3′. Concentrations of H_6Tp20 used were 0, 0.1, 0.3, 1, and 3 μM. Control experiments revealed no binding of 1 μM p24 to ssRNAs. (B) Mobility-shift experiments of H_6Tp20 binding to trace amounts of RNA with a double-stranded section and 5′-single-stranded tail. BS was the 22mer and was 5′-^{32}P-labeled and annealed to excess amounts of unlabeled 12-, 16-, 18-, 20-, and 22mer TS. Formation of the annealed complex was confirmed by a microshift of p*BS upon addition of TS. [Compare (−) and (+), the first and second lanes of each RNA length set, respectively.] Sequences of TAR-derived TS and BS are shown at the top. Concentrations of H_6Tp20 used were as in (A).

first molecule of p24. Subsequent experiments were performed with duplex RNA and ssDNA competitor.

Discrete-length double-stranded oligonucleotides were prepared to test directly the RNA length requirement for binding. These RNAs are derived from TAR sequences and designed to force a single base-pairing register (Figure 3A). A variety of native-gel and filter-binding conditions gave no binding of p24 to dsRNA of 6–16 base pairs, including conditions that give successful binding with longer RNAs. Moreover, binding was not observed in competition experiments in which a p*dsTAR−p24 complex was challenged with 50 μM 8 and 16 base-pair dsRNA (data not shown).

The minimal dsRNAs that bound protein were 16 base pairs for H_6Tp20 (Figure 3A) and 18 bp for p24 (data not shown). H_6Tp20 binding to 16–20 base-pair dsRNA resulted in formation of only complex 1, with complex 2

appearing for 22 and 24 base-pair dsRNA at high protein concentration. This suggests that the minimal-length site for saturation of longer RNAs is 11 base pairs (=22/2), or one turn of A-form dsRNA.

The ability of single-stranded tails to rescue binding of short double-stranded helices was also examined. As an example of the notation used, 12mer top-strand binding to the 5′-^{32}P-labeled 22mer bottom strand is called 12/22. Constructs have 5′-single-stranded overhangs. Very weak binding of H_6Tp20 to 12/22 and weak binding to 16/22 and 18/22 were observed (Figure 3B). Strong binding required 20 base pairs in 20/22. This result suggests that the dsRBD does not strongly interact with single-stranded tails, although a slight dsRNA length rescue is observed. In summary, the binding of the dsRBD from PKR requires a minimum of 16−18 base pairs of dsRNA, is not strongly rescued by single-stranded tails, and is weakened by RNAs with bulges and by tRNA competitor. In addition, the longer p24 construct shows evidence of protein−protein interaction in the presence of dsRNA.

Requirement of 2′-Hydroxyls for dsRBD Binding to dsRNA. In order to assess the role of the 2′-OH, it was first necessary to establish whether the dsRBD from PKR could bind to RNA−DNA hybrids. Mobility shifts for RNA−DNA hybrids were examined under conditions that give band shifts with an RNA−RNA duplex of identical sequence. RNA−DNA hybrids and dsDNA did not support band shifts with H_6Tp20 (Figure 4A) or p24 (data not shown), indicating that hybrids cannot bind as well as dsRNA.

It was possible, however, that hybrids could not support mobility shifts but could bind weakly to the protein. If so, hybrids should be able to compete with limiting amounts of radiolabeled dsRNA for binding to polypeptide. As shown in Figure 4B, neither dsDNA or RNA−DNA or DNA−RNA hybrids, at concentrations to 100 μM, competed effectively with trace amounts of 5′-^{32}P-labeled dsRNA for binding to H_6Tp20. Only unlabeled dsRNA itself was able to compete with release of 5′-^{32}P-labeled dsRNA. The inability of hybrids to compete was not affected by use of different buffer conditions (Figure 4B; see Discussion). The ability of the dsRBD from PKR to discriminate against RNA−DNA duplexes suggests a direct role for the 2′-OH on both strands of dsRNA in recognition of the dsRBD from PKR.

To look more closely at the 2′-OH requirement for binding, a series of chimeric duplexes was designed and their ability to bind to H_6Tp20 was tested. A 22 base-pair duplex was substituted with 2′-H or 2′-OCH_3 in 12 of 44 sugars in three different orientations: on the same face of the duplex one turn of the helix apart (SF substituted), clustered in the middle of the duplex (MID substituted), and on opposite faces of the duplex one and one-half turns apart (OF substituted) (Figure 5A). Consider first results for 2′-deoxy substitutions. Binding was strongest for the OF-substituted duplex with a K_d of 0.3 μM, compared to 0.2 μM for the unsubstituted duplex (Figure 5B). Binding to SF- and MID-substituted duplexes was somewhat weaker with K_ds of 0.6 and 2 μM, respectively. The MID-substituted duplex gave rise to two band shifts as with the unsubstituted duplex, while the SF- and OF-substituted duplexes gave primarily a single band (Figure 5B). Similar results were obtained with 2′-OCH_3-substituted chimeras (data not shown), with the OF-substituted duplex again binding tightest. Curiously, whereas MID-2′-deoxy-substituted oligomers led to two band shifts,

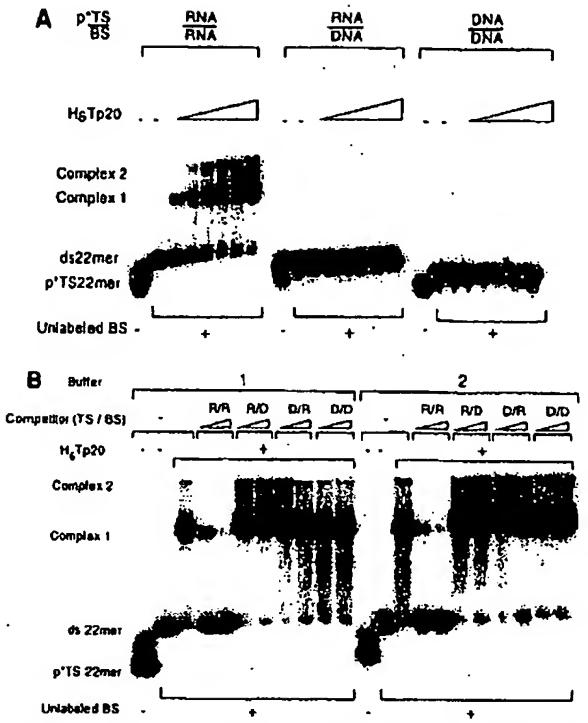


FIGURE 4: No binding of RNA−DNA hybrids or of dsDNA. (A) Mobility-shift experiment for H_6Tp20 binding to trace amounts of 22mer double-stranded nucleic acids. Top-strand (TS) oligomer was 5′-^{32}P-labeled and annealed to excess amounts of unlabeled bottom-strand (BS) oligomer. Formation of duplex was confirmed by a microshift of p*TS upon addition of BS. [Compare (−) and (+), the first and second lanes of each RNA length set, respectively.] Sequences of nucleic acids are (TS) 5′CUGGGUUCCCUGGUUUCGGUCU3′ and (BS) 5′AGACCGAAACCAGGGAACCCAG3′; rU was replaced by dU in all-deoxy strands. Concentrations of H_6Tp20 used were 0, 0.06, 0.2, 0.6, 2, 6, and 18 μM. Mobility shifts were detected only for RNA−RNA duplexes, with formation of two complexes. (B) Competition experiments for H_6Tp20 (3 μM) binding to trace amounts of 22 base-pair p*dsRNA; sequence of dsRNA as in Figure 3A. Formation of duplex was confirmed as described above. A no-competitor shift is shown in the third lane of each set. Protein was added to a mixture of trace 22 base-pair p*dsRNA and 10 or 100 μM double-stranded competitor with indicated TS/BS combinations; R = RNA and D = DNA. DNA strands are with rU replaced by dT. [Replacement of rU by rT has little effect on the activity of PKR (Baglioni et al., 1981), suggesting the difference between U and T is not significant for binding.] Buffer 1 is the 1×BB containing 25 mM Hepes (pH 7.5), 10 mM NaCl, 5% glycerol, 5 mM DTT, 0.1 mM EDTA, and 0.1 mg/mL herring sperm DNA; and buffer 2 contains 10 mM Tris (pH 8.0), 10 mM NaCl, 10% glycerol, 0.5 mM DTT, 25 mM KCl, 1 mM $MgCl_2$, 0.2 mM ATP, and 0.1 mg/mL BSA (Bass et al., 1994).

MID-2′-OCH_3-substituted oligomers led to only a single band shift.

The relative mobilities of duplexes on native gels provide information about their conformation (Bhattacharyya et al., 1990; Roberts & Crothers, 1992). Nonchimeric duplexes ran in the anticipated order dsDNA > RNA−DNA hybrid > dsRNA, and all 2′-H and 2′-OCH_3 chimeric duplexes ran similarly to each other and to dsRNA (Figure 5C). Similar mobilities of chimeric duplexes and dsRNA suggest that these duplexes have similar conformations. Thus, results with chimeric substitutions likely reflect atomic interactions and not differences in helical conformation (see Discussion).

Chemical Footprinting of the dsRBD−dsRNA Complex. To determine whether H_6Tp20 protects the minor groove of dsRNA, Fe(II) EDTA chemical footprinting experiments

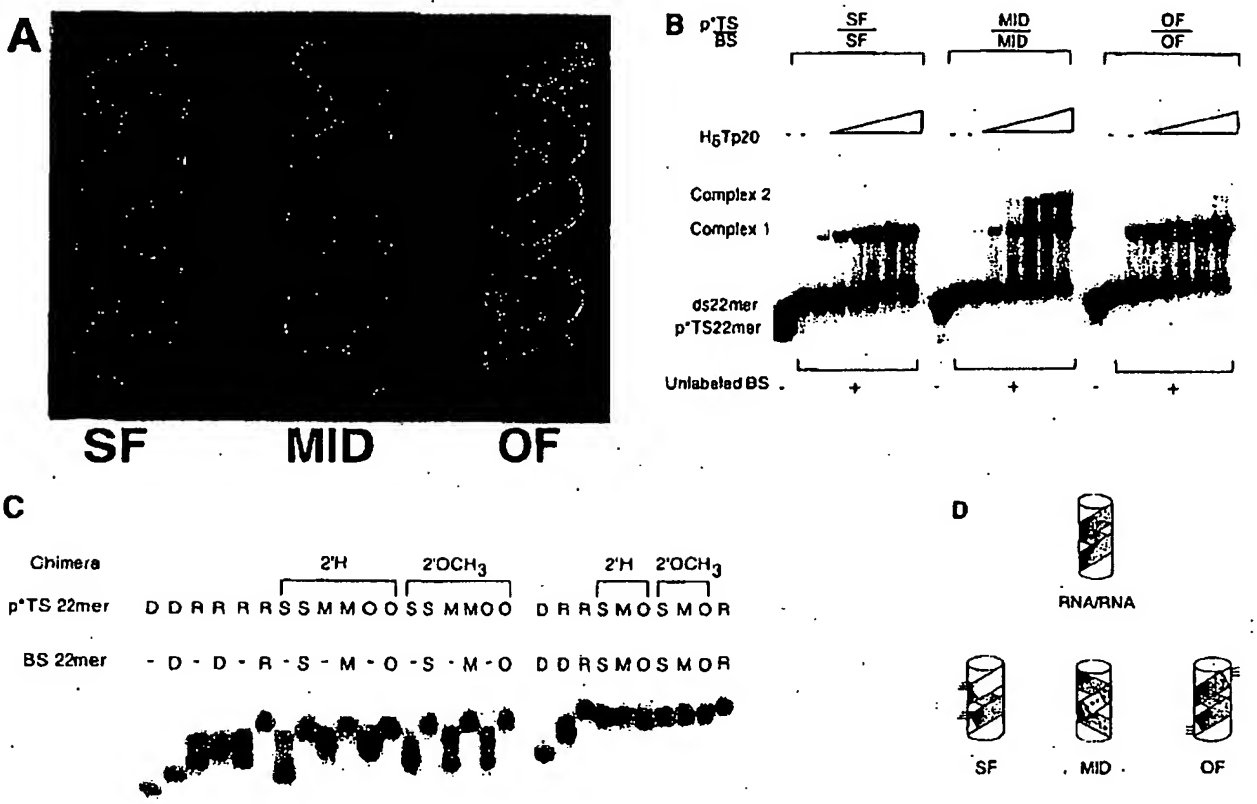


FIGURE 5: Binding of chimeras. (A) Computer-generated views of A-form 22 base-pair duplexes. Green balls show positions of 2′-deoxy substitutions. Note that the 2′-OHs are located in the wide, shallow, and accessible minor groove of A-form dsRNA. Twenty-two base pairs give two full helical turns. In each duplex, 12 of a possible 44 2′-OHs were substituted. SF = 2′-OH substitution in two sets of six on the same face (SF) of the duplex, shown on the left; MID = 2′-OH substitutions in the middle (MID) of the duplex, shown in the center; OF = 2′-OH substitutions in two sets of six on opposite faces (OF) of the duplex, shown on the right. Positions of 2′-OH substitutions for SF substitutions are in italics; MID substitutions are in lower case; and OF substitutions are underlined: top strand (TS), 5′CUGGG*UUC*-ccuggu*UUCGGUCU*3′; bottom strand (BS), 5′A*GAC*C*GAA*accaggG*AAC*C*CAG*3′. 2′-rU is substituted with 2′-dU or 2′-OCH_3U. (B) Native-gel experiment for H_6Tp20 binding to trace amounts of duplex. Experimental conditions were the same as in Figure 4A. K_ds are 0.2 μM for RNA−RNA, 0.6 μM for SF−SF, 2 μM for MID−MID, and 0.3 μM for OF−OF. (C) Comparison of native-gel mobility of various dsRNA, 2′-H- and 2′-OCH_3-containing chimeric duplexes, and RNA−DNA hybrids. Gel conditions were the same as described in Materials and Methods for mobility-shift experiments. The left-hand portion of the gel shows confirmation of duplex formation by a microshift of p*TS upon addition of BS. (Compare the first and second lanes of each duplex set.) The right-hand portion of the gel shows relative mobility of duplexes with the RNA−RNA duplex loaded twice to provide a reference line. D = DNA, R = RNA, S = same face substituted chimeric strand, M = middle substituted chimeric strand, and O = opposite face substituted chimeric strand. (D) Model of H_6Tp20 contact on chimeric duplexes. The cylinder represents 22 bp, or two helical turns, of A-form dsRNA. The diagonal stripes represent the minor groove of the helix, the shaded stripes represent regions of contact with H_6Tp20, and each dash represents *two* deoxy sugars. The unsubstituted duplex is the minimal length of dsRNA that can accommodate two H_6Tp20s; thus its entire minor groove is shaded. The SF-substituted duplex data are consistent with the existence of one unperturbed site with 1.5 deoxy base pairs at each end. The MID-substituted duplex data are consistent with two suboptimal sites, one at each end of the duplex. The OF-substituted duplex data are consistent with six optimal sites in the center of the duplex but not with the binding of two H_6Tp20s as for the unsubstituted duplex.

were performed. Free radicals (presumably OH•) generated by solvent-based Fe(II) EDTA have been useful for probing DNA structure and RNA secondary and tertiary structure in a sequence-independent manner (Hertzberg & Dervan, 1984; Tullius & Dombroski, 1986; Latham & Cech, 1989; Celander & Cech, 1990, 1991; Murphy & Cech, 1993). In particular, the probe is thought to react with the sugar moiety of the backbone to afford strand scission (Hertzberg & Dervan, 1984; Tullius & Dombroski, 1986). Experiments on tRNA suggest the probe reports on the accessibility of the ribose 1′- and 4′-hydrogens (Latham & Cech, 1989), located in the minor groove of an A-form RNA helix. Experiments were designed with a duplex region that has chimeric OF substitutions to allow near wild-type binding and help limit the number of registers on the duplex sampled by the polypeptide. In addition, an eight-nucleotide 5′-single-stranded tail was present in some of the experiments to serve as an internal control for OH• cleavage. Single-stranded and double-stranded regions have been shown to have similar reactivity to OH• cleavage (Celander & Cech, 1990).

Experiments were performed with excess protein and limiting concentrations of ^{32}P-labeled duplex. Control experiments in which the RNA−protein complex was treated with cleavage reagents and then run on a native gel showed complete band shifts of nucleic acid to a single complex, identical to mobility shifts with untreated complex (data not shown). This suggests that the RNA−protein complex is stable to the cleavage conditions used.

As shown in Figure 6A, the double-stranded region was protected by H_6Tp20 from cleavage by the free-radical probe for both top- and bottom-strand 5′-^{32}P-labeled experiments. Quantitation of these experiments is shown in Figure 6B. In

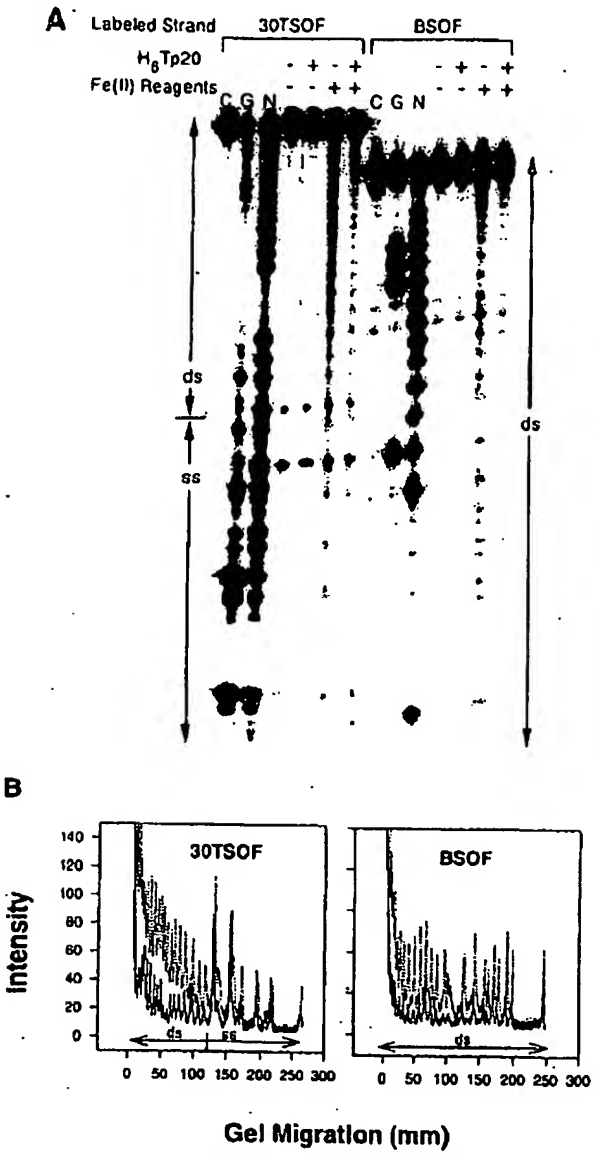


FIGURE 6: Fe(II) EDTA mapping. Fe(II) EDTA footprinting of an annealed top-strand (TS) 30mer−bottom strand (BS) 22mer complex with an eight nucleotide 5′-single-stranded end and 22 base-pair core. The core duplex is chimeric with 2′-H substitutions in the opposite face (OF) orientation. Sequences are as follows, with positions of deoxy substitution underlined: TS, 5′GGAGUGCGCUGGGUUCCCUGGUUUCGGUCU3′; BS, 5′AGACCGAAACCAGGGAACCCAG3′. (A) Denaturing 25% gel showing Fe(II) EDTA mapping. A trace amount of 5′-^{32}P-labeled 30TSOF was annealed to excess BSOF (left-hand portion of the gel), and a trace amount of 5′-^{32}P-labeled BSOF was annealed to excess 30TSOF (right-hand portion of the gel). In indicated lanes, H_6Tp20 was added at 6 μM (enough to give complete mobility shift of the complex), and in indicated lanes Fe(II) reagents were added. G, C, and N are RNase T1, control T1, and alkaline digests, respectively, of the labeled strand only. Double-stranded (ds) and single-stranded (ss) regions are marked. (B) Intensity versus gel migration for the final two Fe(II) reagent-treated lanes of each radiolabeled oligomer set in (A). Minus-protein lane is represented by a dotted line (···) and plus-protein lane by a solid line (—). An equal number of cpm of radioactivity were loaded in each lane. The loading of equal amounts of radioactivity in each lane was confirmed by integration of the PhosphorImager scans.

top-strand-radiolabeled experiments, H_6Tp20 reduced the cleavage of the double-stranded region by 60%, while cleavage of the single-stranded region was reduced by only

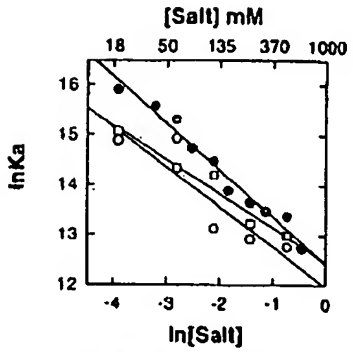


FIGURE 7: Salt dependence. Dependence of the natural logarithm of the association constant on the natural logarithm of the monovalent salt concentration for H_6Tp20 binding to 20 base-pair dsRNA; sequence as in Figure 3A. Formation of duplex was confirmed as in Figure 3, and binding to H_6Tp20 gave only a single complex. The slope gives 1.05 contacts for NaCl (●) (Record et al., 1976). Similar slopes are obtained for a NaOAc (○) and KCl (□) corresponding to 1.05 and 0.8 ion pairs, respectively.

20%. The apparent 20% protection of the single-stranded region could be due to nonspecific association of the protein with the single-stranded tail, although other effects such as quenching of free radicals by the protein could cause apparent protection. However, the 40% difference in cleavage between the double- and single-stranded regions of the RNA can be assigned to preferential protection of the double-stranded RNA by the protein. In bottom-strand-radiolabeled experiments, H_6Tp20 protected the bottom strand to a similar extent (50%). In both experiments, protection of the double-stranded region is fairly uniform, suggesting that much of the minor groove is protected by polypeptide (Figure 6B).

Determination of the Number of Ion Pairs between the dsRBD and dsRNA. Record and co-workers (1976) developed a quantitative theory that describes the number of ion pairs formed between protein and nucleic acid in terms of release of thermodynamically bound monovalent cations from the nucleic acid. A plot of ln K_a versus ln [salt] yields estimates of both the electrostatic and nonelectrostatic components of binding free energy (Record et al., 1976; Lohman et al., 1980). The slope, m, of the plot is related to the number of ion pairs, Z, between the phosphate backbone and protein by $m = -Z\Psi$, where Ψ is the fraction of counterion thermodynamically bound per phosphate. Ψ is equal to 0.89 for poly(A)·poly(U) (Record et al., 1976), and this value was used as an estimate of Ψ for the 20mer dsRNA used here.

In order to look at RNA−protein and not protein−protein interactions, binding of H_6Tp20 to 20mer dsRNA, which gives a single band shift even at high protein concentration, was studied. In addition, since divalent metal is not required for binding, it was omitted from these experiments in order to simplify the interpretation of the data. The slope for NaCl-dependence experiments is 0.94, corresponding to 1.05 (=0.94/0.89) ion pairs (Figure 7). Replacement of either the cation by K^+ or the anion by OAc^- resulted in similar dependencies (Figure 7), consistent with a general ion effect rather than an effect of specific association of either the cation or anion with the protein. Extrapolation of the fit in Figure 7 to 1 M NaCl (the y-intercept) allows calculation of the nonelectrostatic contribution to binding (Record et al., 1976; Lohman et al., 1980; Witherell & Uhlenbeck, 1989). Estimating that the ion pair destabilizes binding by 0.2 kcal/mol at 1 M NaCl (Record et al., 1976; Lohman et al., 1980),

the nonelectrostatic component of binding is −7.6 kcal/mol, accounting for approximately 90% of the total free energy of binding at physiological salt concentration.

DISCUSSION

The dsRBM is an evolutionarily conserved module which enables diverse proteins to bind dsRNA (St Johnston et al., 1992; Kharrat et al., 1995). Two recent NMR analyses of single copies of the motif revealed a structurally compact domain (Bycroft et al., 1995a; Kharrat et al., 1995). The dsRBM binds dsRNA in a sequence-independent manner (Hunter et al., 1975), suggesting that RNA recognition by the dsRBM is unique with respect to known RNP complexes. We find that the dsRBD from PKR binds dsRNA but not RNA−DNA or DNA−RNA hybrids. Our data suggest this discrimination exists because the dsRBD makes only one ion pair with the phosphate backbone, which is similar between dsRNA and hybrids, and instead largely relies on a series of nonelectrostatic 2′-OH interactions throughout the binding site involving both strands of dsRNA.

Two dsRBMs of PKR Facilitate Strong Binding of dsRNA. Native-gel and filter-binding experiments with a series of C-terminal truncated polypeptides indicate that two copies of the dsRBM from PKR are needed for strong, dsRNA-specific binding. This result contrasts with reports that polypeptides derived from PKR containing amino acids 1−91 or 1−98, having a full copy of only dsRBM1, bind to dsRNA (McCormack et al., 1992, 1994; Schmedt et al., 1995). In addition, other polypeptides containing only one copy of the dsRBM can fold into stable structures and bind dsRNA, including the third dsRBM from the *Drosophila* staufen protein, the second dsRBM from the *Xenopus* Xlrbpa protein, and the dsRBM from the *Escherichia coli* RNase III protein (St Johnston et al., 1992; Bycroft et al., 1995a,b; Kharrat et al., 1995). In the above cases, however, the polypeptide was either fused to a larger protein, complexed with an antibody, or present at high concentrations that may stabilize the protein. In addition, the dsRBM1 1−91 polypeptide binds roughly 100-fold more weakly than a polypeptide containing both dsRBMs (Schmedt et al., 1995). Requirement of tandem dsRBMs for optimal dsRNA binding has been reported previously for the *Xenopus* 4F protein (Bass et al., 1994).

dsRBD Binding Requires a Minimum of 16 Base Pairs of dsRNA. Data obtained here indicate that H_6Tp20 requires a minimum of 16 base pairs of dsRNA for strong binding to a single site on dsRNA (Figure 3A), and this requirement is not alleviated by a single-stranded tail (Figure 3B). Site-saturation experiments with H_6Tp20 indicate that two polypeptides can bind to 22 or 24 base-pair dsRNA. Ignoring looping of the RNA, overlap of protein binding sites, and dangling protein, this suggests that a single H_6Tp20 occupies a roughly 11 base-pair site on dsRNA, equivalent to one turn of A-form dsRNA (Saenger, 1984). This observation is consistent with studies of p20 binding to a variety of longer discrete-length dsRNAs that showed that, at saturating concentrations of p20, 11 base pairs are the minimal site required for binding (Manche et al., 1992; Schmedt et al., 1995). The observation that the site size for multiple binding (11 bp) is smaller than that for single binding (16 bp) suggests that an adjacent dsRNA-bound protein can compensate for the absence of a longer dsRNA site. Initial results with a dsRBM from another protein, the third dsRBM from *Drosophila*, indicate that the minimal segment of dsRNA needed for binding is also 11 base pairs (Bycroft et al., 1995a).

Experiments with p24 binding to TAR-based oligomers indicate that bulges weaken RNA−protein interaction (Figure 2A, Table 1). Interestingly, PKR's kinase activity is not activated if an average of one mismatch is present every 8 nucleotides in RNA but can be fully activated if the mismatch occurs only once every 45 nucleotides (Minks et al., 1979). In addition, the loop and bulge of TAR are dispensable for inhibition of PKR activation (Gunnery et al., 1992), consistent with a destabilizing effect of bulges.

2′-Hydroxyls of dsRNA Are Involved in Binding. Two functional groups in dsRNA that are accessible for sequence-independent recognition by a protein are the 2′-OH and phosphate. First we will consider data on the 2′-OH. RNA−DNA hybrids, where DNA is either the top or bottom strand, and dsDNA duplexes are unable to bind to dsRBD constructs as assayed both by mobility-shift experiments and by competition experiments including 100 μM competitor duplex (Figure 4). The K_d for the all-RNA version of these hybrids binding to H_6Tp20 is 0.17 μM, and a lower limit of the K_i for the RNA−DNA hybrid is estimated at $\geq 500\ \mu$M ($=5 \times 100\ \mu$M). These dissociation constants lead to a lower limit for the $\Delta\Delta G°$ for discrimination against RNA−DNA hybrids of ≥ 4.7 kcal/mol. Apparently, the dsRBD from PKR recognizes both strands of the dsRNA. Inability of RNA−DNA hybrids to bind to the dsRBD from PKR is consistent with the inability of such hybrids to activate PKR (Hunter et al., 1975; Sen et al., 1978).

The *Xenopus* 4F protein, which contains two tandem copies of the dsRBM and a C-terminal arginine−glycine-rich block, did not support band shifts with RNA−DNA hybrids, but 100mer and 800mer hybrids were able to compete for binding at concentrations of only 50 pM (Bass et al., 1994). This competition, which is in contrast to our results with PKR, cannot be attributed to differences in solution conditions (Figure 4B); it may indicate that structural differences exist among dsRBDs as required by the specific function of the protein or that other RNA-binding motifs within a protein affect its recognition properties. The *Saccharomyces cerevisiae* RNase H protein which has two copies of the dsRBM is able to bind to hybrids; these particular motifs, however, have some variations from the conserved dsRBM (Cerritelli & Crouch, 1995).

Requirement of 2′-OHs for binding was examined further by testing a series of partially 2′-H- and 2′-OCH_3-substituted, chimeric duplexes. The unsubstituted, same-face-substituted (SF), middle-substituted (MID), and opposite-face-substituted (OF) duplexes showed only modest differences in binding (K_ds of 0.2, 0.6, 2, and 0.3 μM, respectively; Figure 5B). The more striking difference in the behavior of these duplexes is in binding stoichiometry. High H_6Tp20 concentrations led to primarily one band shift for the OF-substituted duplex, as opposed to the two band shifts observed for unsubstituted 22 base-pair dsRNA (Figures 4A and 5B). This observation suggests that H_6Tp20 binding is destabilized by deoxyriboses at the end of a binding site, contacts which would be forced on the OF-substituted duplex if it were saturated with two H_6Tp20 molecules (Figure 5D). Likewise, high H_6Tp20 concentrations led to primarily one band shift for the SF-substituted duplex (Figure 5B). This observation suggests that H_6Tp20 binding is also destabilized by deoxyriboses at the center of a binding site, interactions

which would be necessary if the SF-substituted duplex were saturated with two H_6Tp20 molecules (Figure 5D). Together, results with the OF- and SF-chimeric duplexes indicate that 2′-OHs at both the end and middle of the 11 base-pair site contribute to binding.

The destabilization of binding constants for OF- and SF-chimeric duplexes (relative to unsubstituted RNA) is only <2- and <4-fold, respectively. The small magnitude of these changes can be most readily explained by the SF-substituted duplex having one free site for H_6Tp20 binding unaffected by deoxy substitutions and the OF-substituted and unsubstituted duplexes having statistically more unsubstituted free sites. In particular, the observed K_d for binding of the first protein to a nucleic acid with multiple free sites is the K_d for binding to a single site divided by the number of free sites (McGhee & von Hippel, 1974). Observation that 16 bp is the minimal-length dsRNA for binding of a single H_6Tp20 molecule (Figure 3A) suggests that <3 bp flanking both sides of an 11 bp ribose-containing site are needed for binding of the first protein. Given the requirement for 3 base pairs to flank each site, there are six free sites in each of the OF-substituted and unsubstituted duplexes (Figure 5D). These free sites are predicted to reduce the observed K_d for binding of the first protein to OF-substituted and unsubstituted duplexes by 6-fold relative to binding to the SF-substituted duplex, reasonably consistent with the slightly lower K_ds observed.

Binding to MID-substituted molecules led to two band shifts for the 2′-H substituted duplex and one band shift for the 2′-OCH_3-substituted duplex. The smallest contiguous dsRNA site for this molecule is 8 base pairs: there are two of these sites, one at each end of the OF-substituted duplex. Given the minimal site described for the SF-substituted duplex, the MID-substituted duplex has no free sites unaffected by deoxy substitutions with 3 flanking base pairs; there are, however, two suboptimal sites (Figure 5D). The suboptimal nature of the sites explains the 10-fold destabilization in binding. Observation of two band shifts with the 2′-deoxy-MID-substituted duplex for all but the lowest protein concentration suggests that the MID-substituted duplex achieves binding by exploiting cooperative protein−protein interactions, as observed in TAR and dsTAR experiments with the p24 construct (Table 1).

Overall, the binding constants are weaker with methoxy than with deoxy substitutions, and only a single mobility shift was observed. Weakened binding could be due to steric interference of the bulky methoxy group. Data examining PKR activation by a series of 2′-OCH_3-substituted polymeric dsRNAs (rI_n·rC_n) is consistent with these observations. Partially methylated dsRNA (<20% substituted in only one strand) fully activates PKR, while more fully methylated dsRNA (40−100% in only one strand) is unable to activate PKR (Minks et al., 1980). A single mobility shift may arise from the ability of MID 2′-OCH_3-substituted riboses to interact favorably with H_6Tp20 as hydrogen bond acceptors.

dsRNA Binding Specificity Is Not Dominated by Helix Conformation. Comparative native-gel assays report conformational differences between duplexes (Bhattacharyya et al., 1990; Roberts & Crothers, 1992). Our duplexes had relative mobilities as follows: dsDNA > RNA−DNA > 2′-OCH_3 chimeric duplexes ≈ 2′-H chimeric duplexes > RNA−RNA (Figure 5C). Relative mobilities of the non-chimeric duplexes were the same as previously reported (Bhattacharyya et al., 1990; Roberts & Crothers, 1992), indicating that this assay is able to differentiate among an A-form helix (dsRNA), a B-form helix (dsDNA), and an intermediate-form helix for the RNA−DNA hybrid (Salazar et al., 1993). Consistent with native gels reporting helix conformational information, ordering of native-gel mobility is not merely the inverse of molecular weight (i.e., dsDNA < RNA−DNA < 2′-H chimeric duplexes < dsRNA < 2′-OCH_3 chimeric duplexes). Chimeric substrates have mobilities very similar to each other and to dsRNA, suggesting an A-form-like geometry. Solution structure data on chimeric duplexes support this conclusion since the helical properties of the chimeric section of a duplex are closer to A-form than to B-form, and the RNA strand of the chimeric duplex is A-form (Zhu et al., 1995).

Since the 2′-substitutions appear to have little effect on helix geometry, it is likely that effects on binding instead reflect the disruption of atomic interactions. In addition, 10% or 20% ethanol, which can make B-form DNA and chimeric duplexes more A-form-like and rescue RNA conformationally dependent protein binding (Baidya & Uhlenbeck, 1995), had no effect on binding of dsRNA and chimeric duplexes (Bevilacqua and Cech, unpublished results), consistent with the chimera binding data reflecting true atomic interactions and not differences in helical geometry.

Minor Groove of dsRNA Is Protected by Protein. Hydroxyl-radical footprinting experiments indicate that the dsRBD protects the minor groove of dsRNA in a general manner (Figure 6), supporting direct interaction of the dsRBD with the minor groove of dsRNA. Lack of a specific H_6Tp20 footprint, despite the presence of a chimeric background, may be due to some slippage of H_6Tp20 on the chimeric duplex due to the existence of the six overlapping binding sites (previous section). In addition, H_6Tp20 may indirectly block adjacent duplex regions from the Fe(II) EDTA probe by a steric effect. In related experiments with the adenovirus-associated VA RNA, a well-studied RNA hairpin that can inhibit PKR activation (Mathews & Shenk, 1991), three sugars in one strand of the apical stem were protected (Clarke & Mathews, 1995). Thus, in both studies recognition of dsRNA by the dsRBD appears to involve a series of minor-groove 2′-OH interactions.

Minor-groove recognition is observed in the binding of $tRNA^{Ala}$ by its aminoacyl-tRNA synthetase (Musier-Forsyth & Schimmel, 1992). Binding of RNA substrate by a group I catalytic RNA is largely sequence-independent; it involves recognition of a substrate-containing duplex by minor-groove interactions with four 2′-OHs on both strands of the duplex and the exocyclic amine of G in a terminal G·U pair [e.g., see Bevilacqua and Turner (1991), Pyle and Cech (1991), Strobel and Cech (1993, 1995)].

Small Contribution of Phosphates in dsRBD Binding to dsRNA. An experimental approach for determining the number of phosphates bound to protein by ion pairing involves a theory relating the binding constant to the ionic strength (Record et al., 1976). It has been verified experimentally for both RNA− and DNA−protein complexes. Application of this method to the R17 coat protein−RNA hairpin complex indicates 4.8 ion pairs between RNA and protein (Witherell & Uhlenbeck, 1989). The X-ray structure of a very similar RNA−protein complex shows 7 phosphates involved in 11 interactions with the protein, 5 of which involve ion pairs with the basic residues lysine and arginine and 6 of which involve polar interactions with asparagine, serine, or tyrosine, in good agreement with the solution

studies (Valegård et al., 1994) (O. C. Uhlenbeck and H. E. Johansson, personal communication). In addition, a model study involving pentalysine association with DNA indicates the theory accurately describes the number of ion pairs (Lohman et al., 1980).

Studies of specific RNA−protein complexes conclude that tat−TAR binding involves 6 ion pairs (Weeks & Crothers, 1992), R17 coat protein−RNA hairpin binding involves 4−5 ion pairs (Witherell & Uhlenbeck, 1989), U1A RBD−RNA hairpin binding involves at least 5−7 ion pairs (Hall, 1994), and S4−α mRNA binding involves at least 4 ion pairs (Deckman et al., 1987). Considering nonspecific DNA−protein complexes, gene 32 protein binds to native or ssDNA with 2 ion pairs (Jensen et al., 1976), RNase binds to denatured DNA with 7 ion pairs (Jensen & von Hippel, 1976), and *lac* repressor binds to nonspecific DNA with 12 contacts (deHaseth et al., 1977). In sharp contrast, results obtained here indicate only one ion pair in the dsRBD−dsRNA 20 base-pair complex (Figure 7). A substantial number of ionic interactions might make it difficult for a dsRNA-binding protein to discriminate against RNA−DNA hybrids and dsDNA, all of which have similar presentation of their phosphates.

Salt-dependence experiments have suggested that interaction of p20 and PKR with VA RNA involves 5 ion pairs (Clarke et al., 1994). Protection studies of p20 binding to VA RNA indicate 4 phosphates at the base of the apical stem−loop, and 3 phosphates in the complex domain are protected from iodine cleavage (Clarke & Mathews, 1995). These results contrast with observation of a single ion pair between H_6Tp20 and dsRNA observed here. There are a number of potential explanations for this difference: (1) Regions of protection from iodine cleavage may result from solvent exclusion and do not necessarily involve protein−RNA interactions (Schatz et al., 1991; Rudinger et al., 1992). (2) Some of the interactions could be nonionic, as observed in the MS2 protein−RNA complex (Valegård et al., 1994) (O. C. Uhlenbeck and H. E. Johansson, personal communication). (3) p20 may recognize VA RNA differently than dsRNA. (4) Experiments examining the salt dependence of binding to VA RNA examined only one protein concentration, so it is unclear if the data reflect equilibrium binding (Clarke et al., 1994).

Mutagenesis studies on several dsRBDs have provided results consistent with the formation of a single ion pair. Single alanine substitutions in PKR reveal only one of the conserved basic amino acids (K60) as absolutely required for binding by a solid-support poly(I)·poly(C) assay (McMillan et al., 1995), and mutagenesis studies confirm this result (Green & Mathews, 1992; Green et al., 1995). In the case of the third dsRBD from the *Drosophila* staufen protein, mutation of surface residues to alanines identifies one lysine (K50) as absolutely required for binding by a Northwestern assay (Bycroft et al., 1995a). The lysines in these two proteins occupy an equivalent position in the dsRBM-consensus sequence, situated in the loop between the third β-strand and the second α-helix in the α−β−β−β−α secondary structure (Bycroft et al., 1995a), and so may have the same function in dsRNA binding. In these studies, other lysine residues were found to be important but not essential for binding, although K64 in PKR was found to be essential for dsRNA binding in other studies with a Northwestern blot analysis (McCormack et al., 1994; McCormack & Samuel, 1995).

Only dsRBM1 Appears To Contact dsRNA. K60 and K64 are conserved in both dsRBM1 and dsRBM2 (St Johnston et al., 1992); thus, if both dsRBMs were contacting the dsRNA, two ion pairs would be expected. This observation, in connection with the data of Bycroft et al. (1995a) that a single dsRBM from *Drosophila* also requires 11 base pairs of dsRNA, suggests that only one of the two dsRBMs in the dsRBD from PKR is actually contacting dsRNA. Since dsRBM1 appears to be more important than dsRBM2 for dsRNA binding (Green & Mathews, 1992; McCormack et al., 1994; Green et al., 1995; Romano et al., 1995), this suggests that only dsRBM1 directly contacts the minimal-length dsRNAs studied here. Longer dsRNAs are needed to activate full-length PKR, with 33 base pairs the minimal length and 80 base pairs the optimal length (Hunter et al., 1975; Minks et al., 1979; Manche et al., 1992). With these longer RNAs both copies of the dsRBM may contact the dsRNA leading to activation perhaps by a conformational change of the protein. The necessity of dsRBM2 for function in H_6Tp20 binding to short dsRNAs studied here may reflect protein folding requirements.

ACKNOWLEDGMENT

We thank Professor Chuck Samuel and Professor Steve Schultz for providing the plasmid DNA encoding PKR, Professor Steve Schultz and Peter DuCharme for advice on protein expression and purification and for providing purified p24, Anne Gooding and Cheryl Grosshans for oligonucleotide synthesis, Barbara Golden, Hans E. Johansson, and Joanne Bevilacqua for critically reading the manuscript, Kevin Weeks for helpful discussion, Alex Szewczak for help with computer modeling, and Professor Olke C. Uhlenbeck and Hans E. Johansson for discussion on the R17 coat protein−RNA hairpin structure.

REFERENCES

Aboul-ela, F., Kam, J., & Varani, G. (1995) *J. Mol. Biol. 253*, 313−332.
Baglioni, C., Minks, M. A., & De Clercq, E. (1981) *Nucleic Acids Res. 9*, 4939−4950.
Baidya, N., & Uhlenbeck, O. C. (1995) *Biochemistry 34*, 12363−12368.
Bass, B. L., Hurst, S. R., & Singer, J. D. (1994) *Curr. Biol. 4*, 301−314.
Bevilacqua, P. C., & Turner, D. H. (1991) *Biochemistry 30*, 10632−10640.
Bhattacharyya, A., Murchie, A. I. H., & Lilley, D. M. J. (1990) *Nature 343*, 484−487.
Burd, C. G., & Dreyfuss, G. (1994) *Science 265*, 615−621.
Bycroft, M., Grünert, S., Murzin, A. G., Proctor, M., & St Johnston, D. (1995a) *EMBO J. 14*, 3563−3571.
Bycroft, M., Proctor, M., Freund, S. M. V., & St Johnston, D. (1995b) *FEBS Lett. 362*, 333−336.
Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, W. H. Freeman and Co., New York.
Celander, D. W., & Cech, T. R. (1990) *Biochemistry 29*, 1355−1361.
Celander, D. W., & Cech, T. R. (1991) *Science 251*, 401−407.
Cerritelli, S. M., & Crouch, R. J. (1995) *RNA 1*, 246−259.
Clarke, P. A., & Mathews, M. B. (1995) *RNA 1*, 7−20.
Clarke, P. A., Pe'ery, T., Ma, Y., & Mathews, M. (1994) *Nucleic Acids Res. 22*, 4364−4374.
Clemens, M. (1992) *Nature 360*, 210−211.
Deckman, I. C., Draper, D. E., & Thomas, M. S. (1987) *J. Mol. Biol. 196*, 313−322.
deHaseth, P. L., Lohman, T. M., & Record, M. T., Jr. (1977) *Biochemistry 16*, 4783−4790.

Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527−2538.

Feng, G.-S., Chong, K., Kumar, A., & Williams, B. R. G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5447−5451.

Gatignol, A., Buckler-White, A., Berkhout, B., & Jeang, K.-T. (1991) *Science 251*, 1597−1600.

Gatignol, A., Buckler, C., & Jeang, K.-T. (1993) *Mol. Cell. Biol. 13*, 2193−2202.

Gill, S. C., & von Hippel (1989) *Anal. Biochem. 182*, 319−326.

Green, S. R., & Mathews, M. B. (1992) *Genes Dev. 6*, 2478−2490.

Green, S. R., Manche, L., & Mathews, M. B. (1995) *Mol. Cell. Biol. 15*, 358−364.

Gunnery, S., Rice, A. P., Robertson, H. D., & Mathews, M. B. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8687−8691.

Gunnery, S., Green, S. R., & Mathews, M. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11557−11561.

Hall, K. B. (1994) *Biochemistry 33*, 10076−10088.

Hanks, S. K., Quinn, A. M., & Hunter, T. (1988) *Science 241*, 42−52.

Hertzberg, R. P., & Dervan, P. B. (1984) *Biochemistry 23*, 3934−3945.

Hovanessian, A. G. (1993) *Semin. Virol. 4*, 237−245.

Hunter, T., Hunt, T., Jackson, R. J., & Robertson, H. D. (1975) *J. Biol. Chem. 250*, 409−417.

Jensen, D. E., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7198−7214.

Jensen, D. E., Kelly, R. C., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7215−7228.

Katze, M. G., Wambach, M., Wong, M.-L., Garfinkel, M., Meurs, E., Chong, K., Williams, B. R. G., Hovanessian, A. G., & Barber, G. N. (1991) *Mol. Cell. Biol. 11*, 5497−5505.

Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M., & Pastore, A. (1995) *EMBO J. 14*, 3572−3584.

Kim, U., Wang, Y., Sanford, T., Zeng, Y., & Nishikura, K. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11457−11461.

Latham, J. A., & Cech, T. R. (1989) *Science 245*, 276−282.

Lohman, T. M., deHaseth, P. L., & Record, M. T., Jr. (1980) *Biochemistry 19*, 3522−3530.

Maitra, R. K., McMillan, N. A. J., Desai, S., McSwiggen, J., Hovanessian, A. G., Sen, G., Williams, B. R. G., & Silverman, R. H. (1994) *Virology 204*, 823−827.

Manche, L., Green, S. R., Schmedt, C., & Mathews, M. B. (1992) *Mol. Cell. Biol. 12*, 5238−5248.

Mathews, M. B. (1993) *Semin. Virol. 4*, 247−257.

Mathews, M. B., & Shenk, T. (1991) *J. Virol. 65*, 5657−5662.

Mattaj, I. W. (1993) *Cell 73*, 837−840.

McCarthy, J. E. G., & Kollmus, H. (1995) *Trends Biochem. Sci. 20*, 191−197.

McCormack, S. J., & Samuel, C. E. (1995) *Virology 206*, 511−519.

McCormack, S. J., Thomis, D. C., & Samuel, C. E. (1992) *Virology 188*, 47−56.

McCormack, S. J., Ortega, L. G., Doohan, J. P., & Samuel, C. E. (1994) *Virology 198*, 92−99.

McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469−489.

McMillan, N. A. J., Carpick, B. W., Hollis, B., Toone, W. M., Zamanian-Daryoush, M., & Williams, B. R. G. (1995) *J. Biol. Chem. 270*, 2601−2606.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P. H., & Higuchi, M. (1996) *Nature 379*, 460−464.

Milligan, J. F., & Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.

Minks, M. A., West, D. K., Benvin, S., & Baglioni, C. (1979) *J. Biol. Chem. 254*, 10180−10183.

Minks, M. A., West, D. K., Benvin, S., Greene, J. J., Ts'o, P. O. P., & Baglioni, C. (1980) *J. Biol. Chem. 255*, 6403−6407.

Murphy, F. L., & Cech, T. R. (1993) *Biochemistry 32*, 5291−5300.

Musier-Forsyth, K., & Schimmel, P. (1992) *Nature 357*, 513−515.

O'Connell, M. A., Krause, S., Higuchi, M., Hsuan, J. J., Totty, N. F. T., Jenny, A., & Keller, W. (1995) *Mol. Cell. Biol. 15*, 1389−1397.

Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., & Nagai, K. (1994) *Nature 372*, 432−438.

Park, H., Davies, M. V., Langland, J. O., Chang, H.-w., Nam, Y. S., Tartaglia, J., Paoletti, E., Jacobs, B. L., Kaufman, R. J., & Venkatesan, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4713−4717.

Patel, R. C., & Sen, G. C. (1992) *J. Biol. Chem. 267*, 7671−7676.

Polson, A. G., & Bass, B. L. (1994) *EMBO J. 13*, 5701−5711.

Proud, C. G. (1995) *Trends Biochem. Sci. 20*, 241−246.

Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., & Williamson, J. R. (1992) *Science 257*, 76−80.

Puglisi, J. D., Chen, L., Blanchard, S., & Frankel, A. D. (1995) *Science 270*, 1200−1203.

Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628−631.

Record, M. T., Jr., Lohman, T. M., & deHaseth, P. (1976) *J. Mol. Biol. 107*, 145−158.

Roberts, R. W., & Crothers, D. M. (1992) *Science 258*, 1463−1466.

Romano, P. R., Green, S. R., Barber, G. N., Mathews, M. B., & Hinnebusch, A. G. (1995) *Mol. Cell. Biol. 15*, 365−378.

Rould, M. A., Perona, J. J., Söll, D., & Steitz, T. A. (1989) *Science 246*, 1135−1142.

Roy, S., Agy, M., Hovanessian, A. G., Sonenberg, N., & Katze, M. (1991) *J. Virol. 65*, 632−640.

Rudinger, J., Puglisi, J. D., Pütz, J., Schatz, D., Eckstein, F., Florentz, C., & Giegé, R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5882−5886.

Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., & Moras, D. (1991) *Science 252*, 1682−1689.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., & Reid, B. R. (1993) *Biochemistry 32*, 4207−4215.

Samuel, C. E. (1993) *J. Biol. Chem. 268*, 7603−7606.

Schatz, D., Leberman, R., & Eckstein, F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6132−6136.

Schmedt, C., Green, S. R., Manche, L., Taylor, D. R., Ma, Y., & Mathews, M. B. (1995) *J. Mol. Biol. 249*, 29−44.

Schweisguth, D. C., Chelladurai, B. S., Nicholson, A. W., & Moore, P. B. (1994) *Nucleic Acids Res. 22*, 604−612.

Sen, G. C., Taira, H., & Lengyel, P. (1978) *J. Biol. Chem. 253*, 5915−5921.

Steitz, T. A. (1993) in *The RNA World* (Gesteland, R. F., & Atkins, J. F., Eds.) p 219, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

St Johnston, D., Brown, N. H., Gall, J. G., & Jantsch, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10979−10983.

Strobel, S. A., & Cech, T. R. (1993) *Biochemistry 32*, 13593−13604.

Strobel, S. A., & Cech, T. R. (1995) *Science 267*, 675−679.

Thomis, D. C., Doohan, J. P., & Samuel, C. E. (1992) *Virology 188*, 33−46.

Tullius, T. D., & Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469−5473.

Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., & Liljas, L. (1994) *Nature 371*, 623−626.

Weeks, K. M., & Crothers, D. M. (1992) *Biochemistry 31*, 10281−10287.

Weeks, K. M., & Crothers, D. M. (1993) *Science 261*, 1574−1577.

Weeks, K. M., & Cech, T. R. (1995) *Biochemistry 34*, 7728−7738.

Witherell, G. W., & Uhlenbeck, O. C. (1989) *Biochemistry 28*, 71−76.

Wong, I., & Lohman, T. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5428−5432.

Ye, X., Kumar, R. A., & Patel, D. J. (1995) *Chem. Biol. 2*, 827−840.

Zaug, A. J., Dávila-Aponte, J. A., & Cech, T. R. (1994) *Biochemistry 33*, 14935−14947.

Zhu, L., Salazar, M., & Reid, B. R. (1995) *Biochemistry 34*, 2372−2380.

BI9607259

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 273, No. 5, Issue of January 30, pp. 2532–2542, 1998

Identification and Partial Purification of Human Double Strand RNase Activity

A NOVEL TERMINATING MECHANISM FOR OLIGORIBONUCLEOTIDE ANTISENSE DRUGS*

(Received for publication, August 21, 1997, and in revised form, October 30, 1997)

Hongjiang Wu, A. Robert MacLeod‡, Walt F. Lima, and Stanley T. Crooke§

From the Department of Molecular Pharmacology, Isis Pharmaceuticals, Carlsbad, California 92008

We have identified a double strand RNase (dsRNase) activity that can serve as a novel mechanism for chimeric antisense oligonucleotides comprised of 2′-methoxy 5′ and 3′ "wings" on either side of an oligoribonucleotide gap. Antisense molecules targeted to the point mutation in codon 12 of Harvey Ras (Ha-Ras) mRNA resulted in a dose-dependent reduction in Ha-Ras RNA. Reduction in Ha-Ras RNA was dependent on the oligoribonucleotide gap size with the minimum gap size being four nucleotides. An antisense oligonucleotide of the same composition, but containing four mismatches, was inactive.

When chimeric antisense oligonucleotides were prehybridized with 17-mer oligoribonucleotides, extracts prepared from T24 cells, cytosol, and nuclei resulted in cleavage in the oligoribonucleotide gap. Both strands were cleaved. Neither mammalian nor *Escherichia coli* RNase HI cleaved the duplex, nor did single strand nucleases. The dsRNase activity resulted in cleavage products with 5′-phosphate and 3′-hydroxyl termini.

Partial purification of dsRNase from rat liver cytosolic and nuclear fractions was effected. The cytosolic enzyme was purified approximately 165-fold. It has an approximate molecular weight of 50,000–65,000, a pH optimum of approximately 7.0, requires divalent cations, and is inactivated by approximately 300 mM NaCl. It is inactivated by heat, proteinase K, and also by a number of detergents and several organic solvents.

Antisense oligonucleotides have been shown to inhibit gene expression for a number of cellular targets (1). These compounds have proven to be effective research tools and are of interest as therapeutic agents. To date most antisense oligonucleotides studied have been oligodeoxynucleotides. Oligodeoxynucleotides are believed to cause a reduction in target RNA levels through the action of RNase H (2), an endonuclease that cleaves the RNA strand of RNA:DNA duplexes (3). This enzyme, thought to play a role in DNA replication, has been shown to be capable of cleaving the RNA component of oligodeoxynucleotide:RNA duplexes in cell-free systems as well as in *Xenopus* oocytes (4–6). RNase H is very sensitive to structural alterations in antisense oligonucleotides (7), and thus attempts to increase the potency of oligonucleotides by increasing affinity, stability, lipophilicity, and other characteristics by chemical modifications of the oligonucleotide have often resulted in oligonucleotides that no longer generate substrates for RNase H when bound to their target RNA (8). RNase H activity is also somewhat variable (8), thus a given disease state may not be a candidate for antisense therapy simply because the target tissue has insufficient RNase H activity. Therefore it is clear that terminating mechanisms in addition to RNase H are of potential value to the development of antisense therapeutics.

In addition to the pharmacological inhibition of gene expression described above, it is becoming clear that organisms from bacteria to humans use endogenous antisense RNA transcripts to alter the stability of some target mRNAs and regulate gene expression (9, 10). The best characterized cases of antisense-mediated gene regulation are derived from studies on bacteria; for example an endogenous antisense RNA transcript regulates the expression of *mok* mRNA in certain bacteria. As the antisense RNA level drops, *mok* mRNA levels rise, which leads to the induction of a cytotoxic protein (*hok*), resulting in cell death (11). Other systems regulated by such mechanisms in bacteria include the RNA I-RNA II hybrid of the ColE1 plasmid (12), OOP-cII RNA regulation in bacteriophage λ (13), and the copA-copT hybrids in *Escherichia coli* (14). In *E. coli* the RNA:RNA duplexes formed have been shown to be substrates for regulated degradation by the endoribonuclease RNase III. Duplex-dependent degradation has also been observed in the archaebacterium, *Halobacterium salinarium*, where an antisense transcript reduces expression of the early (T1) transcript of the phage gene *phiH* (15).

In bacteria, RNase III is the double strand endoribonuclease responsible for the degradation of some antisense:sense RNA duplexes. RNase III carries out site-specific cleavage of double strand RNA (dsRNA)[1]-containing structures and also plays an important role in mRNA processing and in the processing of rRNA precursors into 16, 23, and 5 S ribosomal RNA (16). In eukaryotes, a yeast gene (*RNT1*) has recently been cloned that codes for a protein that has striking homology to *E. coli* RNase III and shows dsRNase activity as well as a role in ribosomal RNA processing (17). Avian cells treated with interferon produce and secrete a soluble nuclease capable of degrading dsRNA (18); however, such a secreted dsRNase activity is not a likely candidate to be involved in the intracellular degradation of antisense:sense RNA duplexes. Despite these findings, little is known about human or mammalian dsRNase activities.

In this work we have designed chimeric antisense oligonucleotides that contain 2′-methoxy-modified nucleotides in the "wings" and ribonucleotides in the "gap." These compounds bind to their cellular targets with high affinity to form an oligonucleotide:mRNA duplex in cells. Designing a series of

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Current address: MethylGene, 7220 Frederick Banting, Montreal, Quebec H4S 2A1, Canada.

§ To whom correspondence should be addressed: Dept. of Molecular Pharmacology, Isis Pharmaceuticals, 2292 Faraday Ave., Carlsbad, CA 92008. Tel.: 760-603-2311; Fax: 760-931-0265.

[1] The abbreviations used are: ds, double strand; Ha-Ras, Harvey RAS; pCp, cytidine biophosphate.

oligonucleotides with varying ribonucleotide content enabled us to identify, and partially purify, an activity in human cells and rat liver that requires the formation of a dsRNA region (oligoribonucleotide:mRNA) to degrade target RNA in cells. The finding that human cells and rat liver contain an activity capable of recognizing and cleaving dsRNA suggests that human cells may have conserved mechanisms for regulation of gene expression by antisense RNA present in prokaryotes. Further, this activity presents a novel terminating mechanism for antisense drugs. Strategies aiming to exploit this activity to its fullest may have important implications for antisense therapeutics.

MATERIALS AND METHODS

Oligonucleotide Synthesis—RNA gap mer 2'-methoxyphosphorothioate oligonucleotides were synthesized using an Applied Biosystems 380 B automated DNA synthesizer as described previously (19). Oligonucleotides were synthesized using the automated synthesizer and 5'-dimethoxytrityl 2'-tert-butyldimethylsilyl 3'-*O*-phosphoramidite for the RNA portion and 5'-dimethoxytrityl 2'-*O*-methyl 3'-*O*-phosphroamidite for 5' and 3' wings. The protecting groups on the exocyclic amines were phenoxyacetyl for ribodenosine and riboguanosine, benzoyl for ribocytosine and 2'-*O*-methyl A and C, and isobutyl for 2'-*O*-methyl G. The standard synthesis cycle was modified by increasing the wait step after the delivery of tetrazole and base to 600 s repeated four times for RNA and twice for 2'-methoxy. The fully protected oligonucleotide was cleaved from the support, and the phosphate group was deprotected in 3:1 ammonia/ethanol at room temperature overnight, then lyophilized to dryness. Treatment in methanolic ammonia for 24 h at room temperature was then done to deprotect all bases, and the sample was again lyophilized to dryness. The pellet was resuspended in 1 M tetrabutylammonium fluoride in tetrahydrofuran for 24 h at room temperature to deprotect the 2' positions. The reaction was then quenched with 1 M triethylaminoacetate, and the sample was then reduced to 0.6 volume by rotovac before being desalted on a G25 size exclusion column (Boehringer Mannheim). The oligonucleotide recovered was then analyzed spectrophotometrically at 260 nm for yield. Purity was characterized by capillary electrophoresis and by mass spectrometry. In all cases the purity was in excess of 90%.

^{32}P Labeling of Oligonucleotides—The sense oligonucleotide was 5'-end-labeled with ^{32}P using [γ-^{32}P]ATP, T4 polynucleotide kinase, and standard procedures (20). The labeled oligonucleotide was purified by electrophoresis on 12% denaturing polyacrylamide gel electrophoresis (20). The specific activity of the labeled oligonucleotide was approximately 5000 cpm/fmol.

Cell Culture and Northern Blot Analysis—T24 human bladder carcinoma cells were maintained as monolayers in McCoys medium (Life Technologies, Inc.) supplemented with 10% fetal bovine serum and 100 units/ml penicillin. After treatment with oligonucleotide (see below for details) for 24 h, cells were trypsinized and centrifuged, and total cellular RNA was isolated according to standard protocols (20). To quantitate the relative abundance of Ha-Ras mRNA, total RNA (10 μg) was transferred by Northern blotting onto a Bio-Rad Zeta probe membrane (Bio-Rad) and UV cross-linked (Stratalinker, Stratagene, La Jolla, CA). Membrane-bound RNA was hybridized to a ^{32}P-labeled 0.9-kilobase pair Ha-Ras cDNA probe (Oncogene Science, Pasadena, CA) and exposed to XAR film (Eastman Kodak Co.). The relative amount of Ha-Ras mRNA was determined by normalizing the Ha-Ras signal to that obtained when the same membrane was stripped and hybridized with a probe for human glyceraldehyde-3-phosphate dehydrogenase (CLONTECH, Palo Alto, CA). Signals from Northern blots were quantified using a PhosphorImager and Imagequant software (Molecular Dynamics, Sunnyvale, CA).

Oligonucleotide Treatment of Cells—Cells growing as a monolayer were washed once with warm phosphate-buffered saline, then Opti-MEM (Life Technologies, Inc.) medium containing Lipofectin (Life Technologies, Inc.) at a concentration of 5 μg/ml per 200 nM of oligonucleotide up to a maximum concentration of 15 mg/ml was added. Oligonucleotides were added and the cells were incubated at 37 °C for 4 h, after which the medium was replaced with full serum medium. After 24 h in the presence of oligonucleotide, the cells were harvested, and RNA was prepared for further analysis.

RNase H Analysis—RNase H analysis was performed using a chemically synthesized 17-base oligoribonucleotide complementary to bases +23 to +40 of activated (codon 12 mutation) Ha-Ras mRNA. 20 nM of the 5'-end-labeled RNA was incubated with a 100-fold molar excess of the various antisense oligonucleotides in a reaction containing 20 mM Tris-Cl, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, and 4 units of RNase inhibitor (Pharmacia Biotech Inc.) in a final volume of 10 μl. Secondary structures in the oligonucleotides were melted out by heating to 95 °C for 5 min, followed by slow cooling to room temperature. Duplex formation was confirmed by the shift in mobility between the single strand end-labeled sense RNA and the annealed duplex on nondenaturing polyacrylamide gels. The resulting duplexes were tested as substrates for digestion by either *E. coli* RNase HI (U. S. Biochemical Corp., Cleveland, OH) or mammalian RNase HI (partially purified from calf thymus). 1 μl of a 1×10^{-4} mg/ml solution of either *E. coli* RNase HI or mammalian RNase HI was added to 10 μl of the duplex reaction and incubated at 37 °C for 30 min, after which the reaction was terminated by the addition of denaturing loading buffer. Reaction products were resolved on a 12% polyacrylamide gel containing 7 M urea and exposed to XAR film (Kodak).

Cell-free in Vitro Nuclease Assays—Duplexes used in the cell-free T24 extract experiments were annealed as described above. After formation of the duplex the reaction was treated with 1 μl of a mixture of RNase T and A (RPAII kit, Ambion, Austin, TX) and incubated for 15 min at 37 °C, to remove any nonduplexed single strand oligonucleotides. The duplex was then gel-purified from a nondenaturing 12% polyacrylamide gel. T24 cell nuclear and cytosolic fractions were isolated as described previously (21). 10 μl of the annealed duplexes were incubated with 20 μg of the T24 nuclear or cytosolic extract at 37 °C. The reaction was terminated by phenol/chloroform extraction and ethanol-precipitated with the addition of 10 μg of tRNA as a carrier. Pellets were resuspended in 10 μl of denaturing loading dye, and products were resolved on 12% denaturing acrylamide gels as described above. ^{32}P-Labeled 17-base RNA was base-hydrolyzed by heating to 95 °C for 10 min in the presence of 50 mM $NaCO_3$, pH 9.0, to generate a molecular weight ladder.

Duplexes for the rat liver extracts were prepared in 30 μl of reaction buffer (20 mM Tris-HCl (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$, 0.1 mM dithiothreitol) containing 10 nM antisense oligonucleotide and 10^5 cpm of ^{32}P-labeled sense oligonucleotide. Reactions were heated at 90 °C for 5 min and incubated at 37 °C for 2 h. The oligonucleotide duplexes were incubated with either unpurified and semipurified extracts at a total protein concentration of 25 μg of unpurified cytosolic extract, 20 μg of unpurified nuclear extract, 1–4 μl (1–4 μg) ion-exchange-purified cytosolic fraction, or 1–4 μl (100–400 ng) ion-exchange and gel filtration-purified cytosolic fractions or ion-exchange-purified nuclear fraction. Digestion reactions were incubated at 37 °C for 0–240 min. Following incubation, 10 μl of each reaction were removed and quenched by addition of denaturing gel loading buffer (5 μl of 8 M urea, 0.25% xylene cyanol FF, 0.25% bromphenol blue). The reactions were heated at 95 °C for 5 min and resolved in a 12% denaturing polyacrylamide gel. To perform nondenaturing gel analysis, 20 μl of the reaction mixture were quenched by adding 2 μl of the native gel loading buffer (50% glycerol, 0.25% bromphenol blue FF). The reactions were resolved in a 12% native polyacrylamide gel containing 44 mM Tris borate and 1 mM $MgCl_2$. Gels were analyzed using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

Determination of 5' and 3' Termini—Nonlabeled duplex was treated with T24 extracts as described previously. Half of this reaction was then treated with calf intestinal phosphatase (Stratagene) while the other half was left untreated. The phosphatase was inactivated by heating to 95 °C, and the reactions were extracted with phenol/chloroform and then precipitated in ethanol with glycogen as a carrier. The precipitates were then treated with T4 polynucleotide kinase (Stratagene) and [γ-^{32}P]ATP (ICN, Irvine, CA). The samples were again extracted by phenol/chloroform and precipitated with ethanol. The products of the reaction were then resolved on a 12% acrylamide gel and visualized by exposure to Kodak XAR film. The 3' terminus of the cleaved duplex was evaluated by the reaction of duplex digestion products with T4 RNA ligase (Stratagene) and [^{32}P]pCp (ICN).

Liver Extraction and Preparation of Nuclear and Cytosolic Fractions—0.5 kg of rat liver was blended (Waring Commercial Blender, Dynamics Co. of America, New Hartford, CT) and homogenized (Polytron homogenizer, Brinkmann) in 5 ml of buffer X (10 mM Hepes, pH 7.5, 25 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM EDTA, 2 M sucrose, 10% glycerol)/g tissue and centrifuged (Beckman centrifuge J2-21M) at 10,000 rpm for 40 min. The supernatant was precipitated with 40% ammonium sulfate (Sigma). All the activity was recovered in the 40% ammonium sulfate precipitate. The pellet was resuspended in buffer A (20 mM Hepes, pH 6.5, 5 mM EDTA, 1 mM dithiothreitol, 0.25 mM phenylmethylsulfonyl fluoride, 0.1 M KCl, 5% glycerol, 0.1% Nonidet P-40, and Triton X-100) and dialyzed to remove ammonium sulfate.

Approximately 40 g of cytosolic extract were obtained from 0.5 kg of liver.

The crude nuclear pellet was resuspended and homogenized in a glass Dounce homogenizer (Tenbroeck Tissue Grinders, Willard, OH) in buffer Y (20 mM Hepes, pH 7.5, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 25% glycerol). The homogenate was centrifuged at 10,000 rpm for 1.5 h. The supernatant was precipitated with 70% ammonium sulfate. The pellet was resuspended and dialyzed in buffer A. Approximately 5 g of nuclear extract were obtained.

Ion-exchange Chromatography—Nuclear and cytosolic extracts in buffer A were centrifuged at 8,000 × *g* for 10 min, and the supernatants were loaded onto Hi-Trap SP ion-exchange (Pharmacia Biotech, Sweden) columns in fast protein liquid chromotography. They were eluted with a linear gradient of NaCl, and samples were collected, directly analyzed for activity, and measured for protein concentration (Bio-Rad).

Gel Filtration High Performance Liquid Chromatography—Active samples from the ion-exchange chromatography were pooled, concentrated by a centrifugal filter device (Millipore Co., Bedford, MA), applied to a TSK G-3000 column (Toso Haas, Montgomeryville, PA) with running buffer A containing 100 mM NaCl. Samples were collected and UV absorption at 280 nM was determined; then they were directly analyzed for activity and measured for protein concentration. Concentrated fractions from the gel filtration chromatography were subjected to 12% SDS-polyacrylamide gel electrophoresis (20).

RESULTS

Chimeric 2'-Methoxy-Oligoribonucleotides (RNA GAP Mer) Mediate Digestion of Target RNA in T24 Cells—In two previous publications, structure-activity analyses of antisense oligonucleotides specific for codon 12 of the Ha-*ras* oncogene containing various 2'-sugar modifications were reported (22, 23). Although the 2'-modified oligonucleotides hybridized with greater affinity to RNA than did unmodified oligodeoxynucleotides, they were completely ineffective in inhibiting Ha-*ras* gene expression (23). The lack of activity observed with these 2'-modified oligonucleotides was attributed to their inability to create duplexes that could serve as substrates for degradation by RNase H when bound to their target RNAs (22). Because 2'-modified, and more specifically, 2'-methoxy oligonucleotides do not result in the nucleolytic degradation of their target mRNA, they provide a unique tool for the identification of novel nucleolytic activities that become activated when structural changes are introduced to fully modified 2'-methoxy antisense oligonucleotides.

In this study we have introduced ribonucleotide stretches of various lengths into the center of 17-base 2'-methoxy oligonucleotides targeting Ha-Ras mRNA, to form 2'-methoxy-ribonucleotide 2'-methoxyphosphorothioate oligonucleotides (RNA gap mers) (see Fig. 1, *A* and *B*, for structures). When hybridized to the cellular mRNA target, the resulting duplex consists of two regions that are not targets for nucleolytic degradation (the 2'-methoxy "wings") and one oligoribonucleotide:RNA duplex region that is potentially a target for a ribonuclease activity that recognizes RNA:RNA duplexes.

T24 human bladder carcinoma cells contain an activating G213T transversion mutation in the Ha-*ras* gene at the codon 12 position (24). Chimeric RNA gap mer antisense oligonucleotides specific for this mutation were transfected into T24 cells growing in culture. After incubation with oligonucleotides for 24 h, cells were harvested, total cytosolic RNA was isolated, and Northern blot analysis for Ha-Ras mRNA levels was performed. As previously observed, fully modified 2'-methoxy oligonucleotides did not support nucleolytic cleavage of target mRNA and therefore did not lead to a reduction in steady state levels of Ha-Ras mRNA, even at the highest concentration tested (Fig. 2*A*, *top panel, full 2'-methoxy*). An RNA gap mer oligonucleotide with only 3 ribonucleotides in the gap was also incapable of inducing nucleolytic cleavage of the target RNA (Fig. 2*A*, *bottom panel, 3 GAP RNA*). However, T24 cells treated

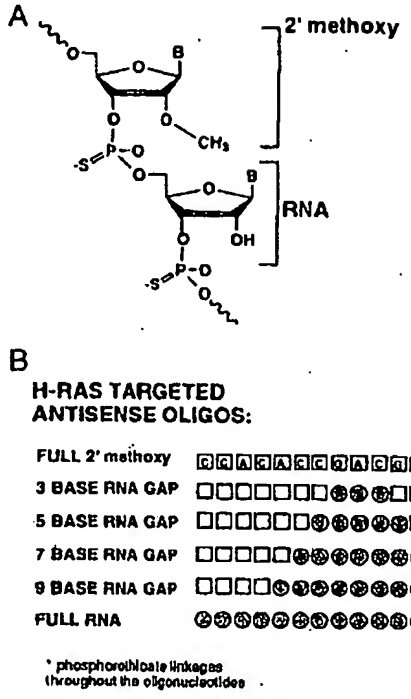


FIG. 1. **Structure of chimeric RNA gap mer oligonucleotides.** *A*, chemical structures show 2 nucleosides of a chimeric 2'-methoxy-ribonucleotide oligonucleotide molecule, with a phosphorothioate linkage between the nucleotides. *B*, schematic shows the design and composition of oligonucleotides used in this study. *Open squares* represent 2'-methoxy-modified nucleotides, *filled circles* represent ribonucleotides. Phosphorothioate linkages are present throughout all the oligonucleotides shown.

with RNA gap mer oligonucleotides containing 5, 7, and 9 ribonucleotides in the gap as well as a full phosphorothioate oligoribonucleotide molecule displayed dose-dependent reductions in Ha-Ras steady state mRNA levels (Fig. 2*B*, *top four panels*, respectively). T24 cells treated with a control 9-base RNA gap mer oligonucleotide that contained four mismatched bases in its sequence did not show dose-dependent reduction in Ha-Ras mRNA suggesting that hybridization to the target RNA was essential for activity (Fig. 2*B*, *bottom panel*). The ability of the RNA gap mer oligonucleotides to reduce Ha-Ras mRNA was dependent on the number of ribonucleotides incorporated into the RNA gap and thus the size of the RNA:RNA duplex formed in cells. The fact that the RNA gap mer oligonucleotide containing three ribonucleotides in the gap was unable to induce reduction in target mRNA suggests that the activity involved requires an RNA:RNA duplex region of at least four ribonucleotides for cleavage of the target. T24 cells treated with 600 nM of the various RNA gap mer oligonucleotides demonstrated a reduction in Ha-Ras mRNA levels of 51 ± 8% for the 5 RNA gap mer, 49 ± 4% for the 7 RNA gap mer, 77 ± 1% for the 9-base RNA gap mer, and 38 ± 5% for the full oligoribonucleotide, respectively, when compared with nontreated controls. The full phosphorothioate oligoribonucleotide molecule was slightly less active than the RNA gap mer oligonucleotides. This is probably due to the relative decrease in stability of the full oligoribonucleotide in cells resulting from inactivation by single stranded ribonucleases, as phosphorothioate 2'-methoxy modified oligonucleotides are considerably more stable than phosphorothioate oligoribonucleotides (25). This suggests that for therapeutic purposes RNA gap mer phosphorothioate oligonucleotides protected by 2'-methoxy wings (or other even more stable 2' modifications) would be more potent molecules. These

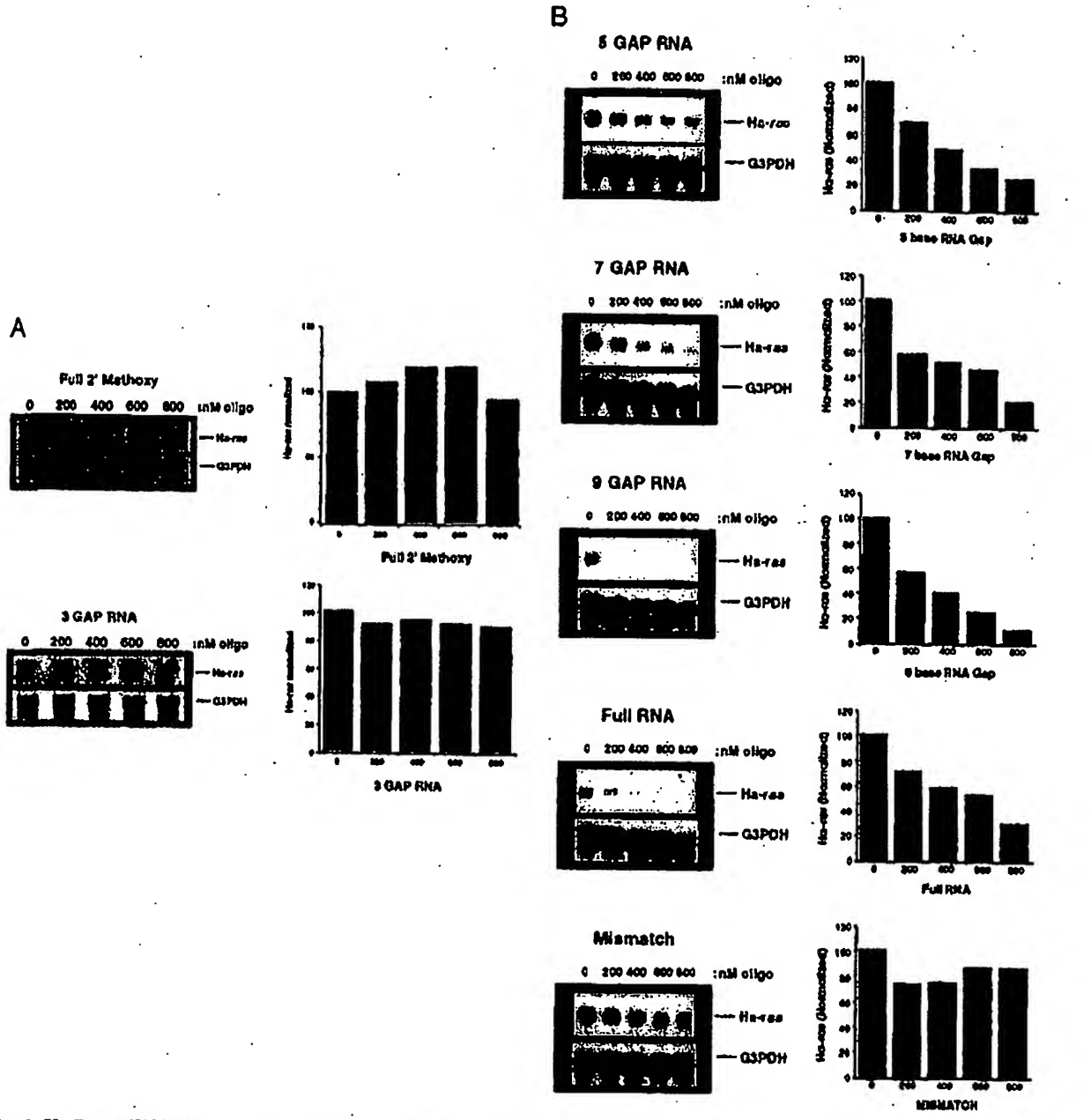


FIG. 2. **Ha-Ras mRNA levels in cells treated with full 2'-methoxy or chimeric RNA gap mer oligonucleotides.** *A*, Northern blot analyses for Ha-Ras mRNA levels in T24 cells treated with the indicated doses of full 2'-methoxy oligonucleotide (*top panel*) or 3-gap oligoribonucleotide (*bottom panel*) for 24 h. The *upper band* is the signal for Ha-Ras. This signal was normalized to that obtained for glyceraldehyde-2-phosphate dehydrogenase (*G3PDH*) (*lower band*), and relative Ha-Ras levels were determined and are presented graphically (*right panel*). Neither oligonucleotide treatment reduced Ha-Ras mRNA levels. *B*, Northern blot analyses of T24 cell treated as in *A*, except with chimeric RNA gap mer oligonucleotides containing either a 5, 7, or 9 ribonucleotide gap or a full ribonucleotide molecule (*top four panels*, respectively). Cells were also treated with a control oligoribonucleotide that contains nine ribose nucleosides with four mismatched bases to the Ha-Ras mRNA sequence (*bottom panel*). Ha-Ras signals were normalized to that of G3PDH, and relative Ha-Ras levels are shown (*right panel*).

experiments demonstrate that an endoribonuclease activity in T24 human bladder carcinoma cells recognizes the internal oligoribonucleotide:RNA portion of a chimeric duplex and reduces target mRNA levels.

An Activity Present in Human Cellular Extracts Induces Cleavage of RNA Gap Mer Oligonucleotide:RNA Duplex within the Internal RNA:RNA Portion in Vitro—To further characterize the dsRNA cleavage activity in T24 cells, we prepared T24 cellular extracts and tested these for the ability to cleave a 17-base pair duplex consisting of the 9-base RNA gap mer oligonucleotide annealed to its complementary ^{32}P-end-labeled oligoribonucleotide. The ^{32}P- labeled duplex was incubated with 20 μg of cytosolic extract at 37 °C for the indicated times (Fig. 3*A*), followed by phenol chloroform extraction, ethanol precipitation, and separation of the products on a denaturing gel. This duplex was a substrate for digestion by an activity present in T24 extracts as can be seen by the loss of full-length end-labeled RNA and the appearance of lower molecular weight digestion products (indicated by *arrows*, Fig. 3*A*). In addition, the activity responsible for the cleavage of the duplex

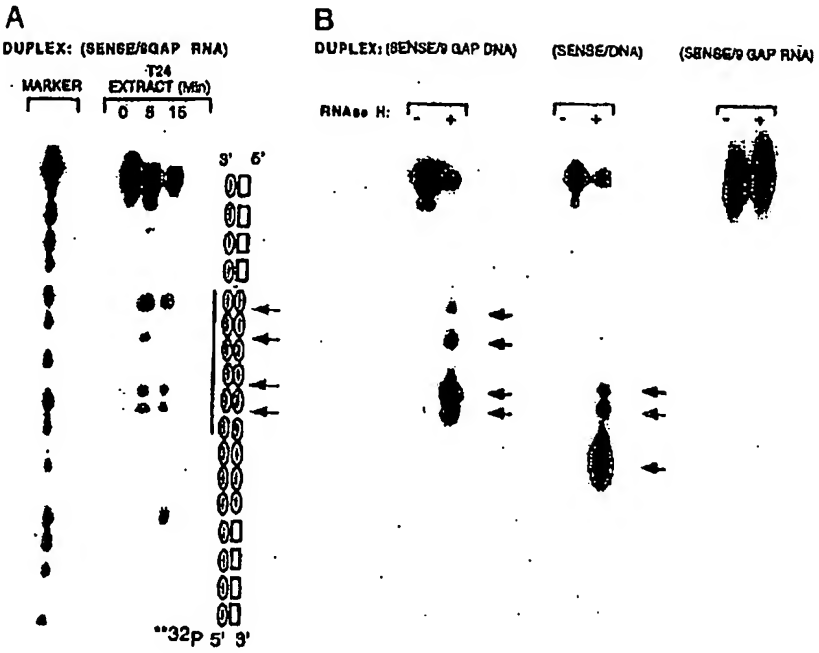


FIG. 3. **Effect of T24 cytosolic extracts and RNase H on duplexes *in vitro*.** *A*, a 17-base pair duplex consisting of the Ha-Ras targeted 9-base RNA gap mer oligonucleotide annealed to a ^{32}P-labeled RNA complement was incubated with 20 μg of a T24 cytosolic protein fraction for the indicated times at 37 °C, the reaction was stopped, and products were resolved on a denaturing polyacrylamide gel. Digestion products (*arrows*) indicate that cleavage of the duplex is restricted to the RNA:RNA region (see schematic of duplex, *far right*). *B*, the same 9-base RNA gap mer oligonucleotide:RNA duplex as in *A* was incubated with or without *E. coli* RNase H (−, +). The lack of digestion products indicates that this duplex is not a substrate for RNase H (*right panel*). Duplexes consisting of ^{32}P-labeled RNA annealed to either a full oligodeoxynucleotide (*middle panel*) or 9-base DNA gap mer oligonucleotide (*left panel*) are substrates for cleavage by RNase H and thus generate digestion products as expected (*arrows*).

displayed specificity for the RNA:RNA portion of the duplex molecule, as indicated by the sizes of the cleavage products it produced (see the physical map of the ^{32}P-end-labeled 9-base RNA gap mer:RNA duplex, Fig. 3*A*, *far right*). To evaluate the cellular distribution of this dsRNase activity, nuclear extracts were prepared from T24 cells and tested for the ability to digest the 9-base RNA gap mer oligonucleotide:RNA duplex. Nuclear extracts prepared from T24 cells were able to degrade the target duplex, and the activity was present in the nuclear fraction at comparable levels to that in the cytoplasmic fractions (data not shown). Cellular extracts prepared from human umbilical vein epithelial cells, human lung carcinoma (A549), and HeLa cell lines all contained an activity able to induce cleavage of the 9-base RNA gap mer:RNA target duplex *in vitro*. This activity was abolished by pretreatment of the extracts with proteinase K for 15 min at 65 °C (data not shown).

The initial RNA gap mer antisense oligonucleotides were synthesized to contain phosphorothioate linkages throughout the entire length of the molecule. As this results in increased stability to single strand nucleases, we reasoned that it would inhibit cleavage of the antisense strand by the dsRNase as well. Therefore, to determine if the activity we have described can cleave both strands in a RNA duplex molecule, we synthesized a 9-base RNA gap mer antisense oligonucleotide that contained phosphorothioate linkages in the wings between the 2′-methoxy nucleotides but had phosphodiester linkages between the nine ribonucleotides in the gap. A duplex comprised of the ^{32}P-labeled 9-base RNA gap mer phosphodiester/phosphorothioate antisense oligonucleotide described above and its complementary oligoribonucleotide was tested as a substrate for the dsRNase activity in T24 extracts. The activity was capable of cleaving the antisense strand of this duplex as well as the sense strand and the pattern of the digestion products indicated that cleavage was again restricted to the RNA:RNA phosphodiester portion of the duplex (data not shown).

An RNA Gap Mer Oligonucleotide:RNA Duplex Is Not a Substrate for RNase HI—To exclude the possibility that the cleavage seen might be due to an RNase H type activity, we tested the ability of *E. coli* RNase H to cleave a 17-base pair duplex composed of the 9-base RNA gap mer oligonucleotide and its complementary 5′-^{32}P-labeled oligoribonucleotide *in vitro*. As can be seen in Fig. 3*B* (*far right panel*), the 9-base RNA gap mer oligonucleotide:RNA duplex was not a substrate for RNase H cleavage as no lower molecular weight bands appeared when it was treated with RNase H. However, as expected both a full oligodeoxynucleotide:RNA duplex and a 9-base DNA gap mer oligonucleotide:RNA duplex were substrates for RNase HI under the same conditions, as is evident by the appearance of lower molecular species in the enzyme-treated lanes (Fig. 3*B*, *left* and *middle panels*). It is interesting to note that RNase HI cleavage of the 9-base DNA gap mer oligonucleotide:RNA duplex (Fig. 3*B*, *left panel*) and cleavage of the 9-base RNA gap mer oligonucleotide:RNA duplex by T24 cellular extracts resulted in similar digestion products (Fig. 3*A*). Both RNase HI and the activity in T24 cells displayed the same preferred cleavage sites on their respective duplexes. Cleavage was restricted to the 3′ end of the target RNA in the region opposite either the DNA or RNA gap of the respective antisense molecule. This suggests that RNase H and the dsRNase activity described here may share binding as well as mechanistic properties.

dsRNase Activity Generates 5′-Phosphate and 3′-Hydroxyl Termini—To determine the nature of the 5′ termini resulting from cleavage of the duplex *in vitro*, nonlabeled duplex was incubated with T24 cellular extracts as described previously, then reacted with T4 polynucleotide kinase and [γ-^{32}P]ATP with or without prior treatment with calf intestinal phosphatase. Phosphatase treatment of the duplex products was essential for the incorporation of the ^{32}P label during the reaction with polynucleotide kinase, indicating the presence of a phosphate group at 5′ termini of digestion products (data not shown). The 3′ termini of the cleaved duplex products were evaluated by the reaction of duplex digestion products with T4 RNA ligase and [^{32}P]pCp. T4 RNA ligase requires a free 3′-hydroxyl terminus for the ligation of [^{32}P]pCp. The ability of the duplex digestion products to incorporate [^{32}P]pCp by T4 RNA ligase indicated the presence of 3′-hydroxyl groups (data not shown).

dsRNase Activity in Rat Liver—To determine if non-human mammalian cells contain dsRNase activity, and to provide a

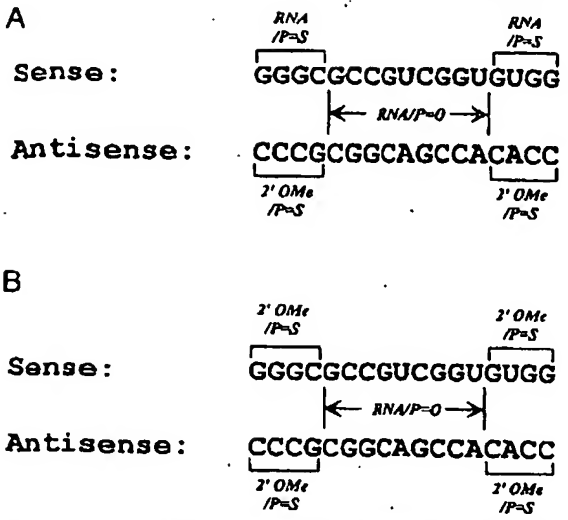


FIG. 4. **Two sets of duplex oligoribonucleotide substrates for the dsRNase activity assay in nondenaturing and denaturing acrylamide gel assays.** *P=O*, phosphodiester linkage; *P=S*, phosphorothioate linkage; *2' Ome*, 2'-methoxy nucleoside. *A*, sense strand has P=S in the wings. *B*, sense strand was 2' Ome and P=S in the wings.

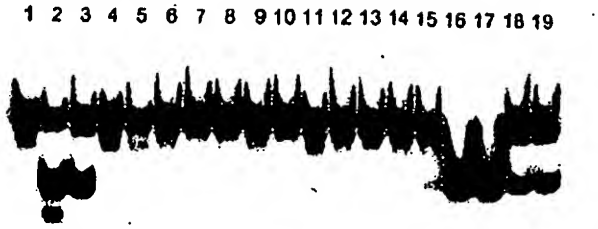


FIG. 5. **Cleavage of substrates by rat liver cytosolic and nuclear extracts.** Antisense and sense oligonucleotides were annealed and incubated with the cellular extracts and variety of RNases, then subjected to native 12% acrylamide gel, as described under "Methods and Materials." *Lane 1*, RNA duplex substrate; *lanes 2* and *3*, duplex digested with partially purified rat liver cytosolic (1 μg) or nuclear extract (0.1 μg); *lane 4*, RNase A (10^{-4} units); *lanes 5* and *6*, RNase CL3 (1 and 10^{-1} unit); *lanes 7* and *8*, partially purified calf thymus RNase H (1/5 and 1/50 unit); *lanes 9* and *10*, *E. coli* RNase H_1 (1/400 and 1/4000 unit); *lanes 11* and *12*: RNase T1 (10^{-1} and 10^{-2} unit); *lanes 13* and *14*, RNase T2 (1 and 10^{-1} unit); *lane 15*, RNase S1 (1 unit); *lanes 16* and *17*, RNase V_1 (1 and 10^{-1} unit); *lanes 18* and *19*, T24 cellular extract (20 and 40 μg).

source from which the activity might be purified, we chose rat liver. In preliminary experiments, dsRNase activity was observed in rat liver homogenates, but the homogenates also displayed higher levels of single strand RNases that confounded analysis because of cleavage of the oligoribonucleotide overhangs after cleavage by dsRNase. To solve this problem, we used two additional substrates and a nondenaturing gel assay. The "antisense" strand in both substrates contained 2'-methoxyphosphorothioate wings on either side of an nine-base ribonucleotide phosphodiester gap. The "sense" strand was either an oligoribonucleotide, with phosphodiester in the 9-base gap flanked by phosphorothioate linkages (Fig. 4*A*), or had flanks comprised of 2'-methoxy nucleosides with phosphorothioate linkages (Fig. 4*B*). Both substrates were more stable to exonuclease digestion than an oligoribonucleotide, and the substrate with phosphorothioate linkages and 2'-methoxy nucleosides in both strands was extremely stable. This was important because of the abundance of single strand RNases relative to the dsRNase activity in the liver and supported the use of nondenaturing assays, as the products of the cleavage by dsRNase remained double-stranded.

Rat liver cytosolic and nuclear extracts induced cleavage of the duplex substrate (Fig. 5, *lanes 2* and *3*). Both extracts resulted in more rapidly migrating bands on native gel electrophoretic analyses. A dsRNase, RNase V1 cleaved the substrate (*lanes 16* and *17*); T24 extracts also cleaved the substrate (*lanes 18* and *19*). Neither bacterial nor human RNase H, nor single-strand RNases cleaved the substrate (*lanes 4–15*).

Fig. 6*A* shows the elution profile of the rat liver cytosolic extract after ion-exchange chromatography. Fig. 6*B* shows that the dsRNase activity eluted in fractions 53–63 (300–450 mM NaCl). In contrast, the dsRNase activity in the nuclear extract eluted at 700–800 mM NaCl (Fig. 6, *C* and *D*). In some chromatographic separations, activities that eluted at both high and low NaCl concentrations were observed in the cytosol and the nucleus.

Fractions from the ion-exchange chromatography of rat liver cytosol were concentrated and subjected to size exclusion chromatography as described under "Materials and Methods." Fig. 7*A* shows the elution profile and Fig. 7*B* the activity profile of cytosolic dsRNase after size-exclusion chromatography. Fig. 7*C* shows a polyacrylamide gel electrophoretic analysis of the concentrated active fractions, after the ion-exchange chromatography, and the fractions from the size exclusion chromatography. The fraction with greatest dsRNase activity (*fraction 3*) had a mean molecular mass of 45–80 kDa, and two bands at approximately 50 kDa appeared to be enhanced on polyacrylamide gel analysis. Comparison of the gel analysis of fractions 3 and 4 shows that proteins of approximately 40 and 64 kDa did not co-purify with the dsRNase activity. *Lane 5* shows that a protein of approximately 55 kDa did not co-purify with the activity. Obviously, fraction 3 represents only a partially purified fraction. Table I provides a summary of the purification and recovery of dsRNase activities from nuclear and cytosolic liver extracts. Purification of the protein(s) responsible for the nuclear activity has proven more difficult and will be the subject of an additional communication.

The effects of various conditions on the dsRNase activity were evaluated using the active fractions after ion-exchange chromatography. Fig. 8 shows that dsRNase activity was apparent in a Tris or phosphate buffer at pH 7–10 (*lanes 1–15*). It was unstable in acetonitrile or methanol (*lanes 42* and *43*) and was inhibited by NaCl; dsRNase activity was inhibited by 30% at 10 mM, >60% at 100 mM, and 100% at 300 mM NaCl (*lanes 36–40*). Heating for 5 min at 60 °C inactivated the enzyme (*lanes 21–23*), and the activity had a temperature optimum of 37–42 °C (*lanes 27–29*). At 25 °C, the activity was approximately 50% of that observed at 37 °C (*lane 30*). The activity was inhibited by EDTA (*lanes 31–35*), required Mg^{2+} and was stable to multiple freeze/thaws (*lanes 24–26*). It also was ablated by treatment with proteinase K (data not shown).

Cleavage Characteristics—To characterize the site of cleavage in more detail, it was necessary to minimize single strand cleavage that occurred after endonuclease cleavage and during handling, particularly after denaturing of the duplex. Consequently, we used the most stable duplex substrate in which both strands of the duplex contained flanking regions comprised of 2'-methoxy nucleosides and phosphorothioate linkages.

Fig. 9*A* displays the results from native gel analyses. *Lane 1* shows the position at which the ^{32}P-labeled sense strand migrated in the native gel. *Lane 2* shows that the "sense" single strand was not digested by dsRNA-specific ribonuclease V1. *Lanes 3* and *4* show the degradation of RNA duplexed with antisense RNA gap mer resulting from high and low concentrations of V1 RNase. *Lanes 5* and *6* show that crude nuclear extract degraded the duplex in a Mg^{2+}-dependent fashion.

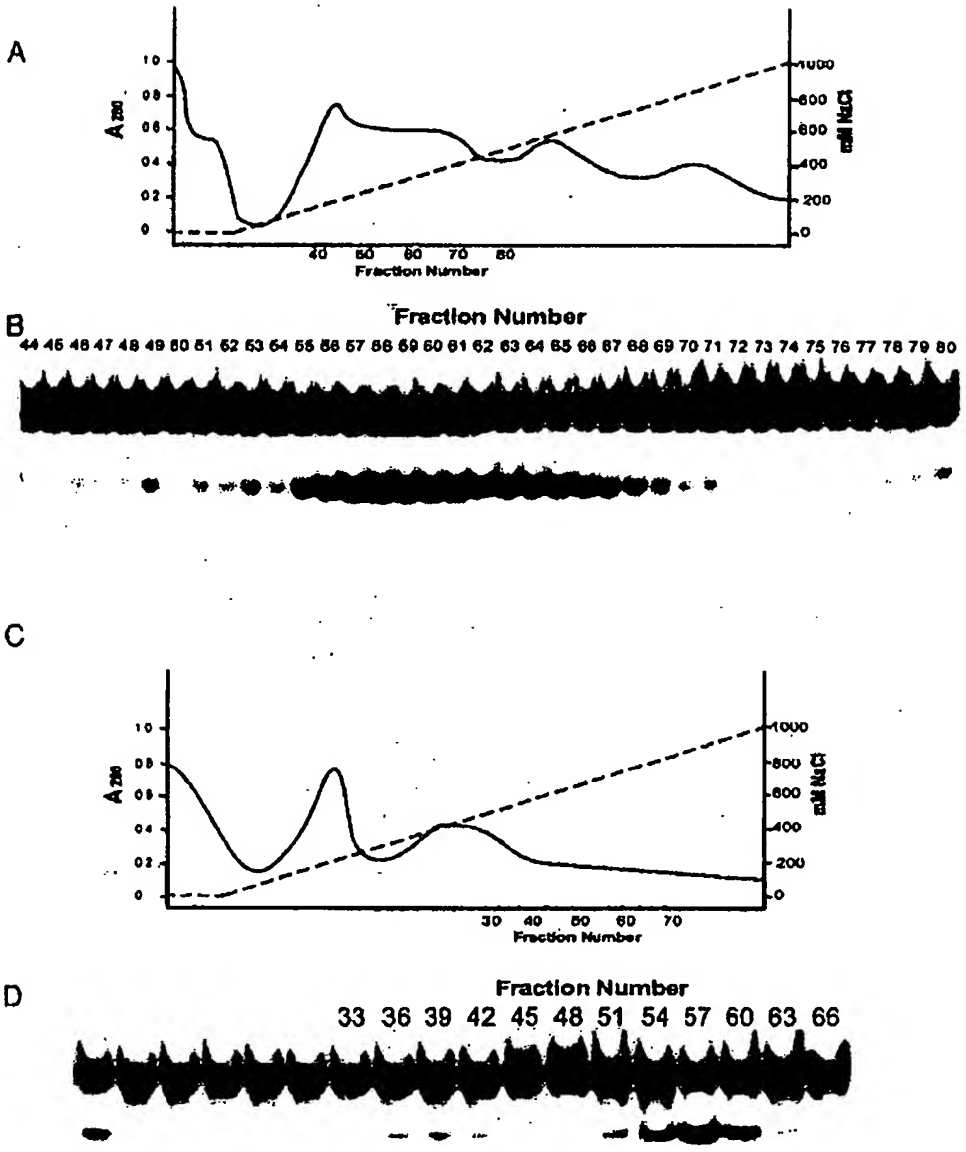


FIG. 6. Ion-exchange chromatograph of dsRNase activity from rat liver cytosolic (*A* and *B*) and nuclear (*C* and *D*) extracts. After NH_4Cl precipitation and dialysis with buffer A, the extracts were loaded onto a 100-ml Hi-Trap SP ion-exchange column and eluted by a 0–1 M NaCl increase gradient (····). *A* and *C*, elution profile; *B* and *D*, dsRNase activity of the fraction (1–2 μl) was determined as described uder "Materials and Methods."

Lane 7 shows that crude cytosolic extract also induced cleavage of the substrate. Ion-exchange purified cytosolic extract cleaved the substrate in a Mg^{2+}-dependent fashion as well (*lanes 8* and *9*). Active fractions alter size exclusion chromatography also cleaved the substrate in a Mg^{2+}-dependent fashion (*lanes 10* and *11*).

Fig. 9*B* shows the denaturing gel analysis of the degradation products. *Lane 1* shows the products of a limit digest of the single-strand sense oligonucleotide. The position of the degradate is consistent with it being the 2′-methoxyphosphorothioate-flanking region (wing). RNase V1 digestion of the single-strand substrate resulted in little degradation (*lane 2*). RNase V1 digestion of the duplex resulted in degradates reflecting cleavage at several sites within the dsRNA gap (*lane 3* and *4*). In *lanes 4–14*, the band at the *top* of the gel demonstrates that, even after denaturation, some of the duplex remained annealed, reflecting the very high affinity of duplexes comprised of 2′-methoxy nucleosides. *Lanes 6–9* show that both the nuclear and cytosolic enzymes cleaved the duplex substrate at several sites within the oligoribonucleotide gap and that the sites of degradation were different from those of V1 nuclease.

DISCUSSION

By the rational design of chemically modified antisense oligonucleotides that contain oligoribonucleotide stretches of varying length, we have identified an activity in cells and rat liver that requires the formation of a dsRNA region to degrade target RNA. This activity is present at comparable levels in both the nuclear and cytoplasmic fractions of T24 human bladder carcinoma cells. We have found that this activity produces 5′-phosphate and 3′-hydroxyl termini after cleavage of its RNA substrate. The generation of 5′-phosphate and 3′-hydroxyl ter-

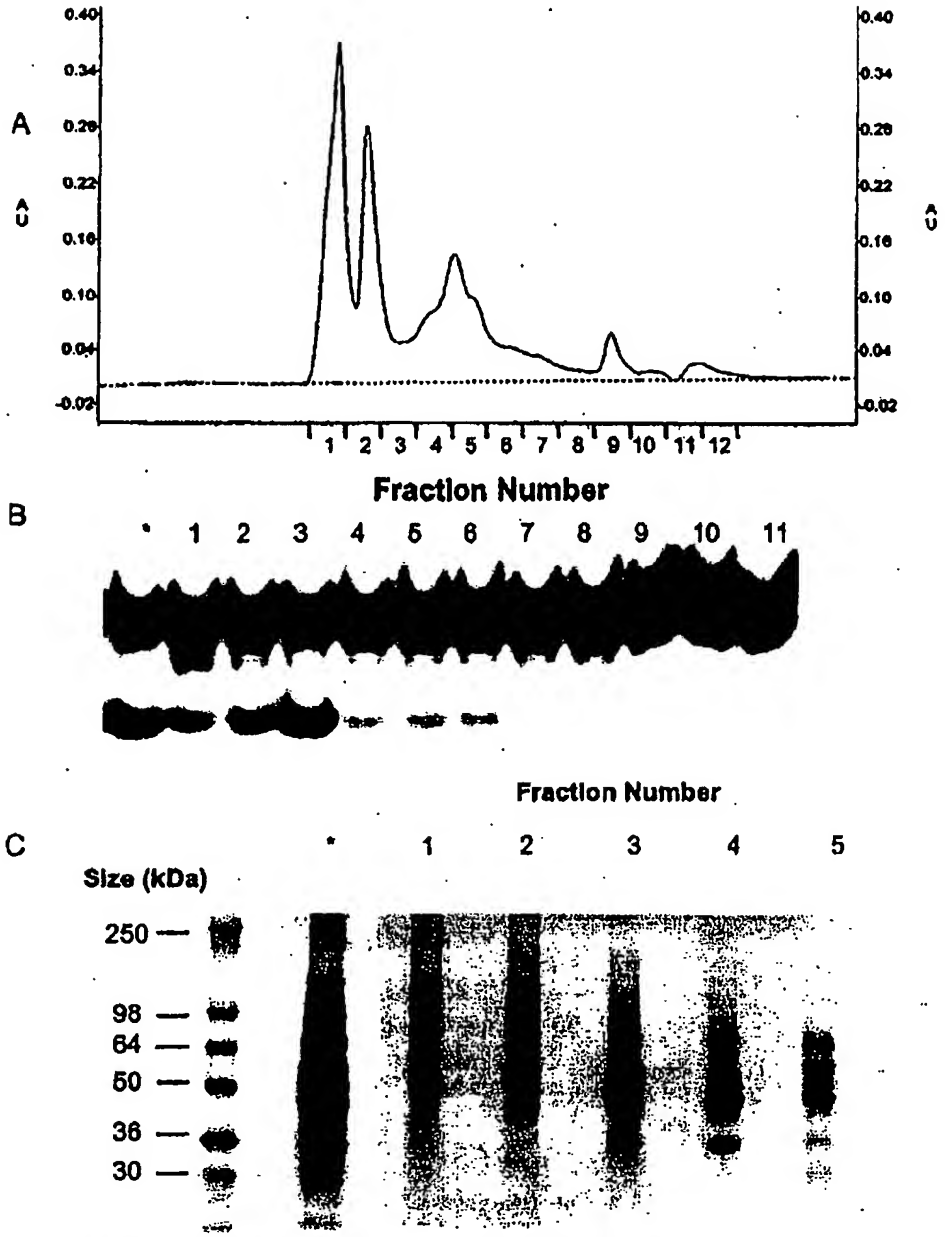


FIG. 7. **Gel filtration of dsRNase activity from rat liver cytosolic extracts.** Extract after ion-exchange was concentrated and loaded onto a TSK3000 gel filtration column. *A*, elution profile; *B*, dsRNase activity for the fractions (1 μl); and *C*, SDS-polyacrylamide gel electrophoresis with Coomassie Blue stain (6 μg of protein from each fraction). • = sample after ion-exchange chromatography only.

TABLE 1
Partial purification of dsRNase from rat liver extracts

Fractions from rat liver nuclei and cytosol were prepared and tested as described under "Materials and Methods."

Fraction	Protein	Total activity	Specific activity	Purification factor	Recovery
	mg	*unit*[a]	*unit/mg*		%
Cytosolic extract	30,000	1,020,000	34	1	100
Ion-exchange pool	991	459,000	463	14	56
Gel filtration pool	18.4	100,980	5,600	165	22
Nuclear extract	5,000	205,000	41	1	100
Ion-exchange pool	11.2	77,900	6970	170	38

[a] Unit is the amount of enzyme required to digest 10 fmol of dsRNA duplex in 15 min at 37 °C in the condition described under "Materials and Methods."

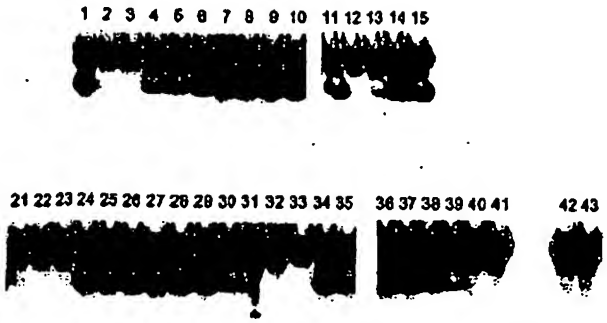


FIG. 8. **Effect of various conditions on dsRNase activity.** 1 μg partially purified rat liver cytosolic extract was incubated with duplex substrate as described under "Materials and Methods." *Lanes 1* and *11*, 20 mM Tris buffer (pH 7.5); *lanes 2–6*, 20 mM sodium acetate buffer (pH 4.5, 5.5, 6.0, 7.0, and 8.0); *lanes 7–10*, 20 mM Tris buffer (pH 7.0, 8.0, 9.0, and 10.0); *lanes 12–15*, 20 mM sodium phosphate buffer (pH 5.0, 6.0, 7.0, and 8.0); *lanes 21–23*, 60, 80, and 100 °C, incubation of extract for 5 min prior to digestion of duplex substrate; *lanes 24–26*, repeat cycles of freezing and thawing 10, 3, and 0 times; *lanes 27–30*, digestion reaction incubated at 50, 42, 37 and 22 °C; *lanes 31–35*, reaction buffer with final EDTA concentration of 50, 20, 10, 5 and 0 mM; *lanes 36–40*, reaction buffer with final NaCl concentration of 30, 100, 300, 500, and 1000 mM; *lane 41*, substrate only; *lanes 42* and *43*, extract pretreated with organic solvent (60% methanol and acetonitrile).

mini is a common feature of several other nucleases that recognize double strand nucleic acid molecules, including RNase HI (26), the enzyme that cleaves the RNA component of a DNA:RNA duplex, and *E. coli* RNase III, which catalyzes the hydrolysis of high molecular weight dsRNA and mediates degradation of sense-antisense duplexes (27). The fact that both the oligoribonucleotide portion of the 9-base RNA gap mer strand in the 9-base RNA gap mer oligonucleotide:RNA duplex as well as the RNA strand were cleaved by this activity demonstrates that the enzyme(s) can specifically recognize and cleave both strands of an RNA:RNA type duplex. The presence of phosphorothioate linkages in the antisense molecule should prevent cleavage of this strand when administered to cells and therefore enhance the potential of such compounds to have therapeutic utility. Interestingly, cleavage of both strands does not seem to be required, in that target mRNA was greatly reduced even though phosphorothioate RNA gap mer antisense oligoribonucleotides were used.

The partial purification of the activity from liver nuclear and cytosolic extracts suggests that the activity is present in both subcellular compartments in rat liver cells as well as human cell lines. The nuclear enzyme eluted from the ion-exchange column at higher NaCl concentrations than did the cytosolic enzymes. However, both require Mg^{2+} and cleave at several sites within the oligoribonucleotide gap. Both require a duplex substrate. This may suggest that there are different types of proteins with dsRNase activity in nuclei and cytosol, but much more work is required before conclusions can be drawn. Additionally, as the nuclear activity eluted at a different NaCl concentration than did the cytosolic, it seems likely that the nuclear activity did not contribute to the cytosolic activity that eluted at lower NaCl concentrations. However, in several preparations, there was evidence of small amounts of activity that eluted at 700–800 mM NaCl in the cytosol, and this could have been due to nuclear contamination. Again, only additional work will definitively determine the cellular localization of the activities.

Many components of mRNA degradation systems have been conserved between pro- and eukaryotes (28, 29). Here we show that like some prokaryotic organisms, in which RNase III carries out the degradation of sense-antisense hybrids to regulate the expression of some genes, human cells have conserved an activity capable of performing a similar role. For some time the dsRNA adenosine deaminase enzyme was suggested to target RNA hybrids for degradation by some unknown mechanism (30). However, more recently it has been demonstrated that deaminated transcripts are usually at least as stable as unmodified RNA (31). This enzyme efficiently modifies duplexes containing 100 base pairs or more and would therefore not be a factor in our system where dsRNA regions ranged from 3 to a maximum of 17 base pairs. In addition, Ha-Ras mRNA does not contain any adenosine residues in the region targeted by our antisense oligonucleotides. The identification of a human dsRNase activity may help us understand how human cells use endogenously expressed antisense transcripts to modulate gene expression. It also has important implications for antisense therapeutics.

The activities reported in this study appear to be novel. The properties of the proteins responsible for cleavage of the substrates are clearly different from other enzymes reported. For example, the dsRNase induced by interferon has a different molecular weight, salt and divalent ion requirements, and is secreted (18). We have not observed dsRNase H activity in cell supernatants.

The vast majority of antisense oligonucleotides used experimentally or currently being tested in the clinic are modified oligodeoxynucleotides (1, 7). It has been demonstrated that the heteroduplex formed between such oligodeoxynucleotide antisense compounds and their target RNA is recognized by the intracellular nuclease RNase H that cleaves only the RNA strand of this duplex. Although RNase H-mediated degradation of target RNA has proven a useful mechanism, it has limitations. One is the fact that the oligonucleotide must be "DNA-like," and such oligonucleotides have inherently a lower affinity for their target RNA. Strategies designed to circumvent this lower affinity include the design of gap mer oligonucleotides that are comprised of a stretch of high affinity chemically modified oligonucleotides on the 5′ and 3′ ends (the wings) with a stretch of deoxynucleotides in the center (the gap) (7, 23). DNA gap mer oligonucleotides have significantly higher affinities for their target. However, depending on the size of the DNA gap, RNase H activity may also be compromised (7, 23). The cellular localization and tissue distribution of RNase H activity are also concerns for antisense therapy. RNase H activity is primarily localized to the nucleus (32), although it has been detected at lower levels in the cytoplasm. RNase H activity is also variable from cell line to cell line and between tissues (8), thus a given disease state may not be a good candidate for antisense therapy, simply because the target tissue has insufficient RNase H activity. Finally, and perhaps most importantly, the majority of sites within RNA targets that have been studied are not sensitive to RNase H-induced cleavage (8). It is clear then that alternative terminating mechanisms to RNase H activation are required for widespread application of antisense therapeutics.

The activity described in this work is attractive as an alternative terminating mechanism to RNase H for antisense therapeutics. The activity relies upon "RNA-like" oligonucleotides that have higher affinity for their target and thus should have higher potency than "DNA-like" oligonucleotides. The presence of the activity in both the cytoplasm and the nucleus suggests that it might be used to inhibit many RNA processing events from nuclear pre-mRNA splicing and transport to the degradation of mature transcripts in the cytoplasm. As we have examined the dsRNase activity induced only by the RNA gap mer oligonucleotides targeted to codon 12 of Ha-Ras, it is difficult to estimate the relative abundance of this dsRNase activity or

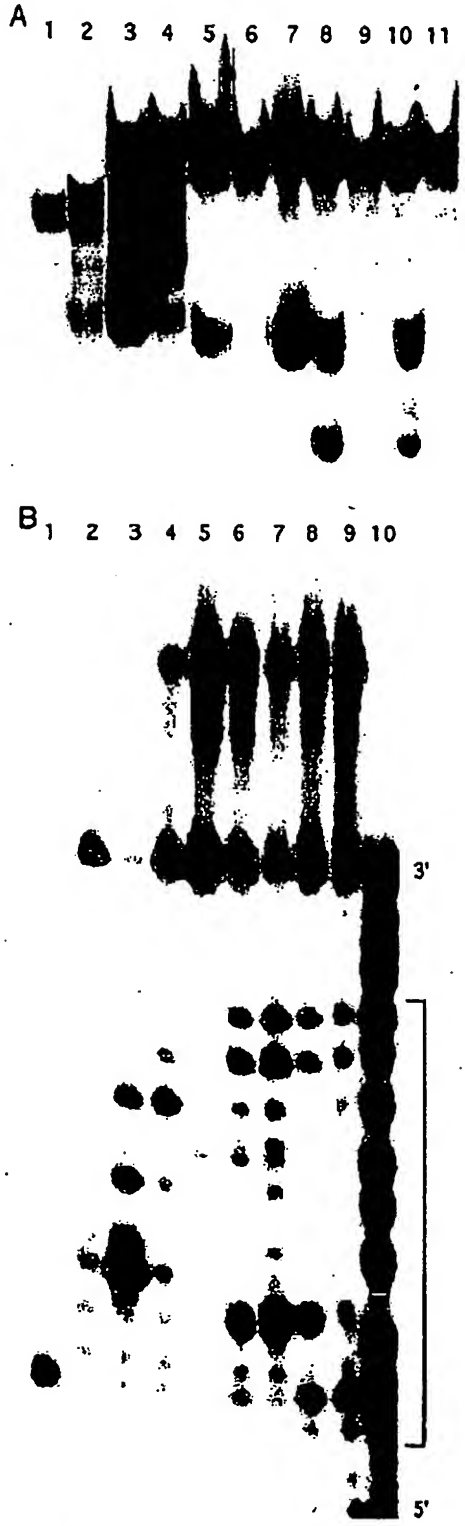


FIG. 9. **Analysis of dsRNA oligonucleotide digestion products by native polyacrylamide gel electrophoresis.** *A*, antisense and sense oligonucleotides were preannealed and incubated with the cellular extracts as described under "Materials and Methods." Polyacrylamide gel analysis of the digestion products was performed as described under "Materials and Methods." Sense strand RNA alone (*lane 1*) and digested with RNase V1 (*lane 2*) are shown. RNase V1 digestion of single strand sense oligonucleotide was performed in 10 μl containing 10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 5 mM $MgCl_2$, 10^4 cpm RNA, and 0.5 unit of RNase V1. RNase V_1 digestion of dsRNA was prepared as above with the exception that 10^4 cpm of sense oligonucleotide was preannealed with 10 nM antisense oligonucleotide prior to digestion with 2×10^{-2} units of RNase V1 (*lane 3*) and 2×10^{-3} units of RNase V1 (*lane 4*). RNase reactions were incubated at 37 °C for 5 min. The digestion patterns for the dsRNA oligonucleotide incubated with the various cellular extracts are as follows: unpurified nuclear extract incubated for 240 min (*lane 5*); unpurified nuclear extract incubated for 240 min in the absence of $MgCl_2$ (*lane 6*); unpurified cytosolic extract incubated for 240 min (*lane 7*); ion-exchange purified cytosolic extract incubated for 240 min (*lane 8*); ion-exchange purified cytosolic extract incubated for 240 min in the absence of $MgCl_2$ (*lane 9*); ion-exchange and gel filtration-purified cytosolic extract incubated for 240 min (*lane 10*); ion-exchange and gel filtration-purified cytosolic extract incubated for 240 min in the absence of $MgCl_2$ (*lane 11*). *B*, analysis of dsRNA oligonucleotide digestion products by denaturing polyacrylamide gel electrophoresis. The *bracketed region* indicates the position of the RNA gap. RNase A and V1 digestions of single strand sense oligonucleotide were performed in 10 μl containing 10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 5 mM $MgCl_2$, 10^4 cpm of ^{32}P-labeled RNA and 5×10^{-4} units of RNase A (*lane 1*) or 2×10^{-2} units of RNase V_1 (*lane 2*). RNase V_1 digestion of dsRNA was performed as described above at 2×10^{-2} units (*lane 3*) or 2×10^{-3} units (*lane 4*). The digestion patterns for the dsRNA oligonucleotide incubated with the various cellular extracts are as follows: unpurified nuclear extract incubated for 0 min (*lane 5*); unpurified nuclear extract incubated for 240 min (*lane 6*); unpurified cytosolic extract incubated for 240 min (*lane 7*); ion-exchange-purified cytosolic extract incubated for 240 min (*lane 8*); ion-exchange and gel filtration-purified cytosolic extract incubated for 240 min (*lane 9*). The base hydrolysis ladder was prepared by incubation of the 10^4 cpm RNA at 90 °C for 5 min in 10 μl containing 100 mM sodium carbonate, pH 9.0 (*lane 10*).

potential potency of these RNA gap mer compounds for other sites compared with RNase H active oligonucleotides. The target site in codon 12 of Ha-Ras is one of the most RNase H-sensitive sites we have identified. A phosphorothioate oligodeoxynucleotide to that site typically displays an IC_{50} of approximately 50 nM in T24 cells (22). The IC_{50} for the 9-base RNA gap mer oligonucleotide was approximately 200 nM, suggesting that this activity is capable of degrading this site nearly as well as RNase H.

The selective inhibition of mutated genes such as the *ras* oncogene necessitates antisense hybridization in the coding region of the mRNA. This requires either a high affinity interaction between oligonucleotide and mRNA to prevent displacement of the oligonucleotide by the polysome or rapid degradation of the target mRNA. RNA gap mer oligonucleotides, being inherently higher in affinity than oligodeoxynucleotides and being able to take advantage of a cellular dsRNase activity, may satisfy both these criteria. Identification of sites that are differentially sensitive to RNase H and to dsRNase activities will increase the number of potential target sites on a given mRNA for antisense oligonucleotides.

It is clear that an activity capable of degrading dsRNA must be carefully regulated, since dsRNA and stem loop structures abound in all cells and uncontrolled cleavage of such substrates would surely be toxic. Mechanisms of regulation may include direct inhibitors and activators, cellular compartmentalization, and regulation by cellular signal transduction pathways. One such pathway that could potentially be involved is the dsRNA-activated protein kinase pathway (33). The kinase p68, which is induced by dsRNA or interferon, phosphorylates the eukaryotic translation initiation factor 2, which results in translational inhibition.

Further purification, characterization, and cloning of the dsRNase activity presented here will be required to increase understanding of its cellular function and regulation. Clearly, the enzyme(s) may play important roles in the intermediary metabolism of RNA and may be involved in the degradation of RNA species targeted by natural antisense transcripts. Drugs

designed to take advantage of this mechanism may help increase the scope of antisense-based therapeutics.

Acknowledgments—We thank P. Villiet for the synthesis of oligonucleotides, F. Bennett, N. Dean, and B. Monia for critical reading of the manuscript and helpful suggestions, and Tracy Reigle for help preparing figures. We also thank Donna Musacchia for excellent administrative assistance.

REFERENCES

1. Crooke, S. T. (1992) *Annu. Rev. Pharmacol. Toxicol.* 32, 329–376
2. Neckers, L., Whitesell, L., Rosolen, A., and Geselowitz, D. A. (1992) *Crit. Rev. Oncol.* 3, 175–231
3. Crouch, R. J., and Dirksen, M. L. (1982) in *Nucleases* (Linn, S. M., and Roberts, R. J., eds) pp. 21–241, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
4. Sankar, S, Cheah, K. C., and Porter, A. G. (1989) *Eur. J. Biochem.* 184, 34–39
5. Dash, P., Lotan, I., Knapp, M., Kandel, E. R., and Goelet, P. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 7896–7900
6. Walter, R. Y., and Walder, J. A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 5011–5015
7. Crooke, S. T., and Bennett, C. F. (1996) *Annu. Rev. Pharmacol.* 36, 107–129
8. Crooke, S. T. (1995) *Therapeutics Application of Oligonucleotides*, pp. 1–25, R. G. Landes Publishers, Austin, TX
9. Nellen, W., and Lichtenstein, C. (1993) *Trends Biochem. Sci.* 18, 419–423
10. Nellen, W., Hildebrandt, M., Mahal, B., Mohrle, A., Kroger, P., Maniak, M., Oberhauser, R., and Sadiq, M. (1992) *Biochem. Soc. Trans.* 20, 750–754
11. Gerdes, K., Nielsen, A., Thorsted, P., and Wagner, E. G. (1992) *J. Mol. Biol.* 226, 637–649
12. Cesareni, G., Helmer-Citterich, M., and Castagnoli, L. (1991) *Trends Genet* 7, 230–235
13. Krinke, L., and Wulff, D. L. (1990) *Genes Dev.* 4, 2223–2233
14. Blomberg, P., Wagner, E. G., and Nordstrom, K. (1990) *EMBO J.* 9, 2331–2340
15. Stolt, P., and Zillig, W. (1993) *Mol. Microbiol.* 7, 875–882
16. Dunn, J. J, and Studier, F. W. (1975) *J. Mol. Biol.* 99, 487–499
17. Elela, S. A., Igel, H., and Ares, M. (1996) *Cell* 85, 115–124
18. Meegon, J., and Marcus, P. I. (1989) *Science* 244, 1089–1091
19. Chiang, M.-Y., Chan, H., Zounes, M. A., Freier, S. M., Lima, W. F., and Bennett, C. F. (1991) *J. Biol. Chem* 266, 18162–18171
20. Ausubel, F. M., Brent, R, Kingston, R. E., Moore, D. D., smith, J. A., Seidman, J. G., and Struhl, K. (1988) *Current Protocols in Molecular Biology*, Chapters 3.10, 4.2, and 10.2, John Wiley & Sons, New York
21. Styf, M., Bozovic, V., and Tanigawa, G. (1991) *J. Biol. Chem.* 266, 10027–10030
22. Monia, B. P., Johnston, J. F., Ecker, D. J., Zounes, M. A, Lima, W. F., and Freier, S. M. (1992) *J. Biol. Chem* 267, 19954–19962
23. Monia, B. P., Lesnik, E. A., Gonzalez, C., Lima, W. F., McGee, D., Guinosso, C. J., Kawasaki, A. M., Cook, P. D., and Freier, S. M. (1993) *J. Biol. Chem.* 268, 14514–14522
24. Weinberg, R. A. (1988) *Cancer* 61, 1963–1972
25. Crooke, S. T., Graham, M. J., Zuckerman, J. E., Brooks, D., Conklin, B. S., Cummins, M. J., Greig, M. J., Guinosso, Kornburst, D., Moncharan, M., Sasmor, H. M., Schleich, T., Tivel, K. L., and Griffey R. H. (1996) *J. Pharmacol. Exp. Ther.* 277, 923–937
26. Crouch, R. J., and Dirksen., M. L. (1982) in *Ribonuclease H in Nucleases* (Linn, S., and Roberts, R. J., eds) pp. 211–241, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Robertson, H. D., Webster, R. E., and Zinder, N. D. (1968) *J. Biol. Chem.* 243, 82–91
28. Schlberg, B., Lundberg, U., Hartl., F. U., and von Gabain, A. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 277–281
29. Wang, M., and Cohen, S. N. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 10591–10595
30. Bass, B. L. (1992) *Semin. Dev. Biol.* 3, 425–433
31. Nishikura, K., Yoo, C., Kim, U., Murry, J. M., Estes, P. A., Cash, F. E., and Linbhaber, S. A. (1991) *EMBO J.* 10, 3523–3532
32. Crum, C., Johnson, J. D., Nelson, A. (1988) *Nucleic Acids Res.* 16, 4569–4572
33. Muers, E. F., Galabru, J., Baber, G. N., Katze, M. G., and Hovanessian, A. G. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 232–236

D3

RECD 10 APR 2003
WIPO PCT

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

April 04, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/386,782*
FILING DATE: *June 06, 2002*
RELATED PCT APPLICATION NUMBER: *PCT/US03/05346*

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

M. SIAS
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

06-10-02

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a req... or filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR §1.53(b)(2).

		Docket No. MBHB02-128-B (900027)	Type a plus sign (+) inside this box:	+

INVENTOR(S)/APPLICANT(S)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (City and either state or foreign country)
Belgelman	Leonid		5530 Coll Dr. Longmont, CO 80503
Macejak	Dennis		6585 Union St. Arvada, CO 80004
Zinnen	Shawn		2378 Birch St. Denver, CO 80207
Pavco	Pamela	A.	708 Barberry Cir. Lafayette, CO 80026

TITLE OF THE INVENTION (280 character maximum)

METHODS AND REAGENTS FOR RNA INTERFERENCE MEDIATED INHIBITION OF GENE EXPRESSION USING SYNTHETIC SHORT INTERFERING RNAS

CORRESPONDENCE ADDRESS

McDonnell Boehnen Hulbert & Berghoff
300 South Wacker Drive, Chicago

STATE	Illinois	ZIP CODE	60606	COUNTRY	U.S.A.

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification — Number of Pages 71 — ☒ Applicant claims benefit of small entity status
☒ Drawing(s) — Number of Sheets 14 — ☒ Other (specify): Certificate of Express Mailing

METHOD OF PAYMENT FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ A check or money order is enclosed to cover the Provisional Filing Fee.	PROVISIONAL APPLICATION FOR PATENT FILING FEE AMOUNT ($)	$80.00
☐ The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number: 13-2490.		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
_X_ No. ____ Yes, the name of the U.S. Government agency and the Government contract number are: ____

Respectfully submitted,

SIGNATURE: _____ Date: June 6, 2002

TYPED or PRINTED NAME Lisa M.W. Hillman REG. NO. 43,673

_X_ Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take 8 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Office of Assistance Quality and Enhancement Division, Patent and Trademark Office, Washington, D.C. 20231, and to the Office of Information and Regulatory Affairs, Office of Management and Budget (Project 0651-0032), Washington, D.C. 20503. DO NOT SEND FEES OR

Express Mail No.
EL 604854251 US

PROVISIONAL APPLICATION FOR PATENT COVER SHEET
Additional Page

		Docket No. MBHB02-128-B (900/027)	Type a plus sign (+) inside this box:	+

INVENTOR(S)/APPLICANT(S)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (City and either state or foreign country)
McSwiggen	James	A.	4668 Franklin Dr. Boulder, CO 80301
Montessy	David		4769 Tanglewood Trail Boulder, CO 80301
Fosnaugh	Kathy		2400 West 17th Avenue 201A Longmont, CO 80501
Mohler	Victor		153 Divide View Dr. Golden, CO 80403
Jamison	Sharon		4985 Twin Lakes Rd #89 Boulder, Co 80301

Express Mail No.
EL 604654251 US

CERTIFICATE OF MAILING

(PATENT)

Express Mail No. EL 604654251US
Deposited June 6, 2002

I hereby certify that the attached correspondence, identified below, is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents, Box Provisional Patent Application, Washington, D.C. 20231.

BY: _____
Lisa M. W. Hillman, Reg. No. 43,673

Application for Provisional Patent of Belgelman et al.

Title: METHODS AND REAGENTS FOR RNA INTERFERENCE MEDIATED INHIBITION OF GENE EXPRESSION USING SYNTHETIC SHORT INTERFERING RNAS

- X — Provisional Patent Application (71 pages of specification and 14 pages of drawings)
- X — Provisional Patent Cover Sheet (2 sheets)
- X — Return Receipt Postcard

Docket No. MBHB02-128-B (900/027)

DESCRIPTION

METHODS AND REAGENTS FOR RNA INTERFERENCE MEDIATED INHIBITION OF GENE EXPRESSION USING CHEMICALLY MODIFIED SYNTHETIC SHORT INTERFERING RNAS

Priority

This application claims the benefit of U.S. Provisional Application 60/358,580, filed February 20, 2002 and U.S. Provisional Application 60/363,214 filed March 11, 2002, both of which are herein incorporated by reference in their entirety, including all drawings.

Background Of The Invention

The present invention concerns methods and reagents useful in modulating gene expression in a variety of applications, including use in therapeutic, diagnostic, target validation, and genomic discovery applications. Specifically, the invention relates to synthetic chemically modified short interfering nucleic acid molecules capable of mediating RNA interference (RNAi).

The following is a discussion of relevant art pertaining to RNAi. The discussion is provided only for understanding of the invention that follows. The summary is not an admission that any of the work described below is prior art to the claimed invention.

RNA interference refers to the process of sequence-specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA) (Fire *et al.*, 1998, *Nature*, 391, 806). The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla (Fire *et al.*, 1999, *Trends Genet.*, 15, 358). Such protection from foreign gene expression may have evolved in response to the production of double stranded RNAs (dsRNA) derived from viral infection or the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA) (Berstein *et al.*, 2001, *Nature*, 409, 363). Short interfering RNAs derived from dicer activity are typically about 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner *et al.*, 2001, *Science*, 293, 834). The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single stranded RNA having sequence complimentary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188).

Short interfering RNA mediated RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. Elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describes RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in Drosophila embryonic lysates (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877) has revealed certain requirements for siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, complete substitution of one or both siRNA strands with 2'-deoxy (2'-H) or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of the 3'-terminal siRNA overhang nucleotides with deoxy nucleotides (2'-H) was shown to be tolerated. Single mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA (Nykanen *et al.*, 2001, *Cell*, 107, 309).

Beach *et al.*, International PCT Publication No. WO 01/68836, describes specific methods for attenuating gene expression using endogenously derived dsRNA. Tuschl *et al.*, International PCT Publication No. WO 01/75164, describes a *Drosophila in vitro* RNAi system and the use of specific siRNA molecules for certain functional genomic and certain therapeutic applications;

although Tuschl, 2001, *Chem. Biochem.*, 2, 239-245, doubts that RNAi can be used to cure genetic diseases or viral infection due "to the danger of activating interferon response". Li *et al.*, International PCT Publication No. WO 00/44914, describes the use of specific dsRNAs for use in attenuating the expression of certain target genes. Zernicka-Goetz *et al.*, International PCT Publication No. WO 01/36646, describes certain methods for inhibiting the expression of particular genes in mammalian cells using certain dsRNA molecules. Fire *et al.*, International PCT Publication No. WO 99/32619, describes particular methods for introducing certain dsRNA molecules into cells for use in inhibiting gene expression. Plaetinck *et al.*, International PCT Publication No. WO 00/01846, describes certain methods for identifying specific genes responsible for conferring a particular phenotype in a cell using specific dsRNA molecules. Mello *et al.*, International PCT Publication No. WO 01/29058, describes the identification of specific genes involved in dsRNA mediated RNAi. Deschamps Depaillette *et al.*, International PCT Publication No. WO 99/07409, describes specific compositions consisting of particular dsRNA molecules combined with certain anti-viral agents. Driscoll *et al.*, International PCT Publication No. WO 01/49844, describes specific DNA constructs for use in facilitating gene silencing in targeted organisms. Parrish *et al.*, 2000, *Molecular Cell*, 6, 1977-1087, describes specific chemically modified siRNA constructs targeting the unc-22 gene of *C. elegans*.

SUMMARY OF THE INVENTION

This invention relates to compounds, compositions, and methods useful for modulating RNA function and/or gene expression in a cell. Specifically, the instant invention features chemically modified synthetic short interfering RNA (siRNA) molecules capable of modulating gene expression in cells by RNA inference (RNAi). The use of chemically modified siRNA is expected to improve various properties of native siRNA molecules through increased resistance to nuclease degradation *in vivo* and/or improved cellular uptake. The chemically modified siRNA molecules of the instant invention provide useful reagents and methods for a variety of therapeutic, diagnostic, agricultural, target validation, genomic discovery, genetic engineering and pharmacogenomic applications.

The introduction of chemically modified nucleotides into nucleic acid molecules will provide a powerful tool in overcoming limitations of *in vivo* stability and bioavailability inherent to native RNA molecules. The use of chemically modified nucleic acid molecules can enable a lower dose of a particular nucleic acid molecule for a given therapeutic effect since chemically modified nucleic acid molecules tend to have a longer half-life in serum. Furthermore, certain chemical modifications can improve the bioavailability of nucleic acid molecules by targeting

particular cells or tissues and/or improving cellular uptake of the nucleic acid molecule. Therefore, even if the activity of a chemically modified nucleic acid molecule is reduced as compared to a native nucleic acid molecule, for example when compared to an all RNA nucleic acid molecule, the overall activity of the modified nucleic acid molecule can be greater than the native molecule due to improved stability and/or delivery of the molecule. Unlike native unmodified siRNA, chemically modified siRNA can also minimize the possibility of activating interferon activity in humans.

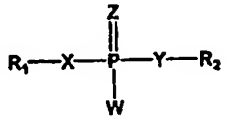
The nucleic acid molecules of the invention that act as mediators of the RNA interference gene silencing response are chemically modified double stranded RNA molecules. As in their native double stranded RNA counterparts, these siRNA molecules typically consist of duplexes containing about 19 base pairs between oligonucleotides comprising about 19 to about 25 nucleotides. The most active siRNA molecules are thought to have such duplexes with overhanging ends of 1-3 nucleotides, for example 21 nucleotide duplexes with 19 base pairs and 2 nucleotide 3'-overhangs. These overhanging segments are readily hydrolyzed by endonucleases *in vivo*. Studies have shown that replacing the 3'-overhanging segments of a 21-mer siRNA duplex having 2 nucleotide 3' overhangs with deoxyribonucleotides does not have an adverse effect on RNAi activity. Replacing up to 4 nucleotides on each end of the siRNA with deoxyribonucleotides has been reported to be well tolerated whereas complete substitution with deoxyribonucleotides results in no RNAi activity (Elbashir et al., 2001, EMBO J., 20, 6877). In addition, Elbashir *et al. supra*, also report that substitution of siRNA with 2'-O-methyl nucleotides completely abolishes RNAi activity. Li *et al.*, International PCT Publication No. WO 00/44914, and Beach *et al.* International PCT Publication No. WO 01/68836 both suggest that siRNA "may include modifications to either the phosphate-sugar back bone or the nucleoside. . . to include at least one of a nitrogen or sulfur heteroatom", however neither application teaches to what extent these modifications are tolerated in siRNA molecules nor provide any examples of such modified siRNA. Kreutzer and Limmer, Canadian Patent Application No. 2,359,180, also describe certain chemical modifications for use in dsRNA constructs in order to counteract activation of double stranded-RNA-dependent protein kinase PKR, specifically 2'-amino or 2'-O-methyl nucleotides, and nucleotides containing a 2'-O or 4'-C methylene bridge. However, Kreutzer and Limmer similarly fail to show to what extent these modifications are tolerated in siRNA molecules nor provide any examples of such modified siRNA.

Parrish *et al.*, 2000, *Molecular Cell*, 6, 1977-1087, tested certain chemical modifications targeting the unc-22 gene in *C. elegans* using long (>25 nt) siRNA transcripts. The authors

describe the introduction of thiophosphate residues into these siRNA transcripts by incorporating thiophosphate nucleotide analogs with T7 and T3 RNA polymerase and observed that "RNAs with two [phosphorothioate] modified bases also had substantial decreases in effectiveness as RNAi triggers (data not shown); [phosphorothioate] modification of more than two residues greatly destabilized the RNAs in vitro and we were not able to assay interference activities." *Id.* at 1081. The authors also tested certain modifications at the 2'-position of the nucleotide sugar in the long siRNA transcripts and observed that substituting deoxynucleotides for ribonucleotides "produced a substantial decrease in interference activity", especially in the case of Uridine to Thymidine and/or Cytidine to deoxy-Cytidine substitutions. *Id.* In addition, the authors tested certain base modifications, including substituting 4-thiouracil, 5-bromouracil, 5-iodouracil, 3-(aminoallyl)uracil for uracil, and inosine for guanosine in sense and antisense strands of the siRNA, and found that whereas 4-thiouracil and 5-bromouracil were all well tolerated, inosine "produced a substantial decrease in interference activity" when incorporated in either strand. Incorporation of 5-iodouracil and 3-(aminoallyl)uracil in the antisense strand resulted in substantial decrease in RNAi activity as well.

Here, applicant discloses the incorporation of various chemical modifications into siRNA constructs. Non-limiting examples of such chemical modifications include without limitation phosphorothioate internucleotide linkages, 2'-O-methyl ribonucleotides, 2'-deoxy-2'-fluoro ribonucleotides, "universal base" nucleotides, 5-C-methyl nucleotides, and inverted deoxyabasic residue incorporation. These chemical modifications, when used in various siRNA constructs, are shown to preserve RNAi activity in cells while at the same time, dramatically increasing the serum stability of these compounds. Furthermore, contrary to the data published by Parrish *et al.*, *supra*, applicant demonstrates that multiple (greater than one) phosphorothioate substitutions are well tolerated and confer substantial increases in serum stability for modified siRNA constructs.

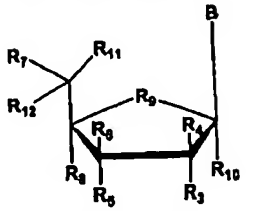
In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises one or more nucleotides comprising a backbone modified internucleotide linkage having Formula I:



wherein each R1 and R2 is independently any nucleotide, non-nucleotide, or polynucleotide which can be naturally occurring or chemically modified, each X and Y is independently O, S, N, alkyl, or substituted alkyl, each Z and W is independently O, S, N, alkyl, substituted alkyl, O-alkyl, S-alkyl, alkaryl, or aralkyl, and wherein W, X, Y and Z are not all O.

The chemically modified internucleotide linkages having Formula I, for example wherein any Z, W, X, and/or Y independently comprises a sulphur atom, can be present in one or both oligonucleotide strands of the siRNA duplex, for example in the sense strand, antisense strand, or both strands. The siRNA molecules of the invention can comprise one or more chemically modified internucleotide linkages having Formula I at the 3'-end, 5'-end, or both 3' and 5'-ends of the sense strand, antisense strand, or both strands. For example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 chemically modified internucleotide linkages having Formula I at the 5'-end of the sense strand, antisense strand, or both strands. In another non-limiting example, an exemplary siRNA molecule of the invention can comprise one or more pyrimidine nucleotides with chemically modified internucleotide linkages having Formula I in the sense strand, antisense strand, or both strands. In yet another non-limiting example, an exemplary siRNA molecule of the invention can comprise one or more purine nucleotides with chemically modified internucleotide linkages having Formula I in the sense strand, antisense strand, or both strands. In another embodiment, a siRNA molecule of the invention having internucleotide linkage(s) of Formula I also comprises a chemically modified nucleotide or non-nucleotide having any of Formulae II, III, V, or VI.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises one or more nucleotides or non-nucleotides having Formula II:

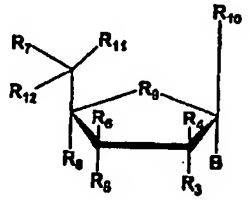


wherein each R3, R4, R5, R6, R7, R8, R10, R11 and R12 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-

alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and B is a nucleosidic base such as adenine, guanine, uracil, cytosine, thymine, 2-aminoadenosine, 5-methylcytosine, 2,6-diaminopurine, or any other non-naturally occurring base that can be employed to form a stable duplex with RNA or a non-nucleosidic base such as phenyl, naphthyl, 3-nitropyrrole, 5-nitroindole, nebularine, pyridone, pyridinone, or any other non-naturally occurring universal base that can be employed to form a stable duplex with RNA.

The chemically modified nucleotide or non-nucleotide of Formula II can be present in one or both oligonucleotide strands of the siRNA duplex, for example in the sense strand, antisense strand, or both strands. The siRNA molecules of the invention can comprise one or more chemically modified nucleotide or non-nucleotide of Formula II at the 3'-end, 5'-end, or both 3' and 5'-ends of the sense strand, antisense strand, or both strands. For example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 chemically modified nucleotide or non-nucleotide of Formula II at the 5'-end of the sense strand, antisense strand, or both strands. In anther non-limiting example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 chemically modified nucleotide or non-nucleotide of Formula II at the 3'-end of the sense strand, antisense strand, or both strands.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises one or more nucleotides or non-nucleotides having Formula III:

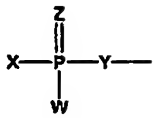


wherein each R3, R4, R5, R6, R7, R8, R10, R11 and R12 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and B is a nucleosidic base such as adenine, guanine, uracil, cytosine, thymine, 2-aminoadenosine, 5-methylcytosine, 2,6-diaminopurine, or any other non-naturally occurring base that can be employed to form a stable duplex with RNA or a non-nucleosidic base such as phenyl, naphthyl, 3-nitropyrrole, 5-nitroindole, nebularine, pyridone, pyridinone, or any other non-naturally occurring universal base that can be employed to form a stable duplex with RNA.

The chemically modified nucleotide or non-nucleotide of Formula III can be present in one or both oligonucleotide strands of the siRNA duplex, for example in the sense strand, antisense strand, or both strands. The siRNA molecules of the invention can comprise one or more chemically modified nucleotide or non-nucleotide of Formula III at the 3'-end, 5'-end, or both 3' and 5'-ends of the sense strand, antisense strand, or both strands. For example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 chemically modified nucleotide or non-nucleotide of Formula III at the 5'-end of the sense strand, antisense strand, or both strands. In anther non-limiting example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 chemically modified nucleotide or non-nucleotide of Formula III at the 3'-end of the sense strand, antisense strand, or both strands.

In another embodiment, a siRNA molecule of the invention comprises a nucleotide having Formula II or III, wherein the nucleotide having Formula II or III is in an inverted configuration. For example, the nucleotide having Formula II or III is connected to the siRNA construct in a 3',3', 3'-2', 2'-3', or 5',5' configuration, such as at the 3'-end, 5'-end, or both 3' and 5' ends of one or both siRNA strands.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises a 5'-terminal phosphate group having Forula IV:



wherein each X and Y is independently O, S, N, alkyl, substituted alkyl, or alkylhalo; each Z and W is independently O, S, N, alkyl, substituted alkyl, O-alkyl, S-alkyl, alkaryl, aralkyl, or alkylhalo; and wherein W, X, Y and Z are not all O.

In one embodiment, the invention features a siRNA molecule having a 5'-terminal phosphate group having Formula IV on the target-complimentary strand wherein the siRNA molecule comprises an all RNA siRNA molecule. In another embodiment, the invention features a siRNA molecule having a 5'-terminal phosphate group having Formula IV on the target-complimentary strand wherein the siRNA molecule also comprises 1-3 nucleotide 3'-overhangs having between about 1 and about 4 deoxyribonucleotides on the 3'-end of one or both strands. In another embodiment, a 5'-terminal phosphate group having Formula IV is present on the target-complimentary strand of a siRNA molecule of the invention, for example a siRNA molecule having chemical modifications having Formula I, Formula II and/or Formula III.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises one or more phosphorothioate internucleotide linkages. For example, in a non-limiting example, the invention features a chemically modified short interfering RNA (siRNA) having about 1, 2, 3, 4, 5, 6, 7, or 8 phosphorothioate internucleotide linkages in one siRNA strand. In yet another embodiment, the invention features a chemically modified short interfering RNA (siRNA) individually having about 1, 2, 3, 4, 5, 6, 7, or 8 phosphorothioate internucleotide linkages in both siRNA strands. The phosphorothioate internucleotide linkages can be present in one or both oligonucleotide strands of the siRNA duplex, for example in the sense strand, antisense strand, or both strands. The siRNA molecules of the invention can comprise one or more phosphorothioate internucleotide linkages at the 3'-end, 5'-end, or both 3' and 5'-ends of the sense strand, antisense strand, or both strands. For example, an exemplary siRNA molecule of the invention can comprise between about 1 and about 5 phosphorothioate internucleotide linkages at the 5'-end of the sense strand, antisense strand, or both strands. In another non-limiting example, an exemplary siRNA molecule of the invention can comprise one or more pyrimidine phosphorothioate internucleotide linkages in the sense strand, antisense strand, or both strands. In yet another non-limiting example, an

exemplary siRNA molecule of the invention can comprise one or more purine phosphorothioate internucleotide linkages in the sense strand, antisense strand, or both strands.

In one embodiment, the invention features a siRNA molecule, wherein the sense strand comprises one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8 , 9 , or 10 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the sense strand; and wherein the antisense strand comprises any of between 1 and 10, specifically about 1, 2, 3, 4, 5, 6, 7, 8 , 9 , or 10 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 pyrimidine nucleotides of the sense and/or antisense siRNA stand are chemically modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3', 5', or both 3' and 5'-ends, being present in the same or different strand.

In another embodiment, the invention features a siRNA molecule, wherein the sense strand comprises between 1 and 5, specifically about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the sense strand; and wherein the antisense strand comprises any of between 1 and 5, specifically about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 pyrimidine nucleotides of the sense and/or antisense siRNA stand are chemically modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without between 1 and 5, for example about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3', 5', or both 3' and 5'-ends, being present in the same or different strand.

In one embodiment, the invention features a siRNA molecule, wherein the antisense strand comprises one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8 , 9 , or 10 phosphorothioate internucleotide linkages, and/or between one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro,

and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the sense strand; and wherein the antisense strand comprises any of between 1 and 10, specifically about 1, 2, 3, 4, 5, 6, 7, 8 , 9 , or 10 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 pyrimidine nucleotides of the sense and/or antisense siRNA stand are chemically modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3', 5', or both 3' and 5'-ends, being present in the same or different strand.

In another embodiment, the invention features a siRNA molecule, wherein the antisense strand comprises between 1 and 5, specifically about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the sense strand; and wherein the antisense strand comprises any of between 1 and 5, specifically about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages, and/or one or more 2'-deoxy, 2'-O-methyl, 2'-deoxy-2'-fluoro, and/or one or more universal base modified nucleotides, and optionally a terminal cap molecule at the 3', 5', or both 3' and 5'-ends of the antisense strand. In another embodiment, one or more, for example about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 pyrimidine nucleotides of the sense and/or antisense siRNA stand are chemically modified with 2'-deoxy, 2'-O-methyl and/or 2'-deoxy-2'-fluoro nucleotides, with or without between 1 and 5, for example about 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages and/or a terminal cap molecule at the 3', 5', or both 3' and 5'-ends, being present in the same or different strand.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule having between about 1 and 5, specifically 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages in each strand of the siRNA molecule.

In another embodiment, a chemically modified siRNA molecule of the invention comprises a duplex having two strands, one or both of which can be chemically modified, wherein each strand is between about 18 and about 27 nucleotides in length, wherein the duplex has between about 18 and about 23 base pairs, and wherein the chemical modification comprises a structure having Formula I, Formula II, Formula III and/or Formula IV. For example, an

exemplary chemically modified siRNA molecule of the invention comprises a duplex having two strands, one or both of which can be chemically modified with a chemical modification having Formula I, Formula II, Formula III, and/or Formula IV, wherein each strand consists of 21 nucleotides, each having 2 nucleotide 3'-overhangs, and wherein the duplex has 19 base pairs.

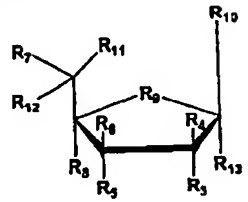
In another embodiment, a siRNA molecule of the invention comprises a single stranded hairpin structure, wherein the siRNA is between about 36 and about 70 nucleotides in length having between about 18 and about 23 base pairs, and wherein the siRNA can include a chemical modification comprising a structure having Formula I, Formula II, Formula III and/or Formula IV. For example, an exemplary chemically modified siRNA molecule of the invention comprises a linear oligonucleotide having between 42 and 50 nucleotides that is chemically modified with a chemical modification having Formula I, Formula II, Formula III, and/or Formula IV, wherein the linear oligonucleotide forms a hairpin structure having 19 base pairs and a 2 nucleotide 3'-overhang.

In another embodiment, a linear hairpin siRNA molecule of the invention contains a stem loop motif, wherein the loop portion of the siRNA molecule is biodegradable. For example, a linear hairpin siRNA molecule of the invention is designed such that degradation of the loop portion of the siRNA molecule *in vivo* can generate a double stranded siRNA molecule with 3'-overhangs, such as 3'-overhangs comprising about 2 nucleotides.

In another embodiment, a siRNA molecule of the invention comprises a circular nucleic acid molecule, wherein the siRNA is between about 38 and about 70 nucleotides in length having between about 18 and about 23 base pairs, and wherein the siRNA can include a chemical modification, which comprises a structure having Formula I, Formula II, Formula III and/or Formula IV. For example, an exemplary chemically modified siRNA molecule of the invention comprises a circular oligonucleotide having between 42 and 50 nucleotides that is chemically modified with a chemical modification having Formula I, Formula II, Formula III, and/or Formula IV, wherein the circular oligonucleotide forms a dumbbell shaped structure having 19 base pairs and 2 loops.

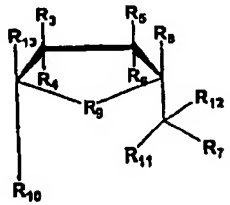
In another embodiment, a circular siRNA molecule of the invention contains two loop motifs, wherein one or both loop portions of the siRNA molecule is biodegradable. For example, a circular siRNA molecule of the invention is designed such that degradation of the loop portions of the siRNA molecule *in vivo* can generate a double stranded siRNA molecule with 3'-overhangs, such as 3'-overhangs comprising about 2 nucleotides.

In one embodiment, a siRNA molecule of the invention comprises at least one abasic residue, for example a compound having Formula V:



wherein each R3, R4, R5, R6, R7, R8, R10, R11, R12, and R13 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having Formula I; R9 is O, S, CH2, S=O, CHF, or CF2.

In one embodiment, a siRNA molecule of the invention comprises at least one inverted abasic residue, for example a compound having Formula VI:



wherein each R3, R4, R5, R6, R7, R8, R10, R11, R12, and R13 is independently H, OH, alkyl, substituted alkyl, alkaryl or aralkyl, F, Cl, Br, CN, CF3, OCF3, OCN, O-alkyl, S-alkyl, N-alkyl, O-alkenyl, S-alkenyl, N-alkenyl, SO-alkyl, alkyl-OSH, alkyl-OH, O-alkyl-OH, O-alkyl-SH, S-alkyl-OH, S-alkyl-SH, alkyl-S-alkyl, alkyl-O-alkyl, ONO2, NO2, N3, NH2, aminoalkyl, aminoacid, aminoacyl, ONH2, O-aminoalkyl, O-aminoacid, O-aminoacyl, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, or group having

Formula I; R9 is O, S, CH2, S=O, CHF, or CF2, and either R2, R3, R8 or R13 serve as points of attachment to the siRNA molecule of the invention.

In another embodiment, a siRNA molecule of the invention comprises an abasic residue having Formula II or III, wherein the abasic residue having Formula II or III is connected to the siRNA construct in a 3',3', 3'-2', 2'-3', or 5',5' configuration, such as at the 3'-end, 5'-end, or both 3' and 5' ends of one or both siRNA strands.

In one embodiment, the invention features a chemically modified short interfering RNA (siRNA) molecule capable of mediating RNA interference (RNAi) inside a cell, wherein the chemical modification comprises a conjugate covalently attached to the siRNA molecule. In another embodiment, the conjugate is covalently attached to the siRNA molecule via a biodegradable linker. In one embodiment, the conjugate molecule is attached at the 3'-end of either the sense strand, antisense strand, or both strands of the siRNA. In another embodiment, the conjugate molecule is attached at the 5'-end of either the sense strand, antisense strand, or both strands of the siRNA. In yet another embodiment, the conjugate molecule is attached both the 3'-end and 5'-end of either the sense strand, antisense strand, or both strands of the siRNA, or any combination thereof. In one embodiment, a conjugate molecule of the invention comprises a molecule that facilitates delivery of a siRNA molecule into a biological system such as a cell. In another embodiment, the conjugate molecule attached to the siRNA is a poly ethylene glycol, human serum albumin, or a ligand for a cellular receptor that can mediate cellular uptake. Examples of specific conjugate molecules contemplated by the instant invention that can be attached to siRNA molecules are described in Vargeese *et al.*, US Serial No. 60/311,865, incorporated by reference herein.

In one embodiment, the invention features a siRNA molecule capable of mediating RNA interference (RNAi) inside a cell, wherein one or both strands of the siRNA comprise ribonucleotides at positions withing the siRNA that are critical for siRNA mediated RNAi in a cell. All other positions within the siRNA can include chemically modified nucleotides and/or non-nucleotides such as nucleotides and or non-nucleotides having Formula I, II, III, IV, V, or VI, or any combination thereof to the extent that the ability of the siRNA molecule to support RNAi activity in a cell is maintained.

In one embodiment, the invention features a method for modulating the expression of a target gene within a cell, comprising: (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of

the target gene; and (b) introducing the chemically modified siRNA molecule into a cell under conditions suitable to modulate the expression of the target gene in the cell.

In one embodiment, the invention features a method for modulating the expression of a target gene within a cell, comprising: (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target gene and wherein the sense strand sequence of the siRNA is identical to the complimentary sequence of the target RNA; and (b) introducing the chemically modified siRNA molecule into a cell under conditions suitable to modulate the expression of the target gene in the cell.

In another embodiment, the invention features a method for modulating the expression of more than one target gene within a cell, comprising: (a) synthesizing chemically modified siRNA molecules of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target genes; and (b) introducing the chemically modified siRNA molecules into a cell under conditions suitable to modulate the expression of the target genes in the cell.

In another embodiment, the invention features a method for modulating the expression of more than one target gene within a cell, comprising: : (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target gene and wherein the sense strand sequence of the siRNA is identical to the complimentary sequence of the target RNA; and (b) introducing the chemically modified siRNA molecules into a cell under conditions suitable to modulate the expression of the target genes in the cell.

In one embodiment, the invention features a method of modulating the expression of a target gene in a tissue explant, comprising: (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target gene; (b) introducing the chemically modified siRNA molecule into a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the target gene in the tissue explant, and (c) optionally introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the target gene in that organism.

In one embodiment, the invention features a method of modulating the expression of a target gene in a tissue explant, comprising: (a) synthesizing a chemically modified siRNA

molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target gene and wherein the sense strand sequence of the siRNA is identical to the complimentary sequence of the target RNA; (b) introducing the chemically modified siRNA molecule into a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the target gene in the tissue explant, and (c) optionally introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the target gene in that organism.

In another embodiment, the invention features a method of modulating the expression of more than one target gene in a tissue explant, comprising: (a) synthesizing chemically modified siRNA molecules of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target genes; (b) introducing the chemically modified siRNA molecules into a cell of the tissue explant derived from a particular organism under conditions suitable to modulate the expression of the target genes in the tissue explant, and (c) optionally introducing the tissue explant back into the organism the tissue was derived from or into another organism under conditions suitable to modulate the expression of the target genes in that organism.

In one embodiment, the invention features a method of modulating the expression of a target gene in an organism, comprising: (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target gene; and (b) introducing the chemically modified siRNA molecule into the organism under conditions suitable to modulate the expression of the target gene in the organism.

In another embodiment, the invention features a method of modulating the expression of more than one target gene in an organism, comprising: (a) synthesizing chemically modified siRNA molecules of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of the target genes; and (b) introducing the chemically modified siRNA molecules into the organism under conditions suitable to modulate the expression of the target genes in the organism.

In one embodiment, the invention features a pharmaceutical composition comprising a chemically modified siRNA molecule of the invention in a pharmaceutically acceptable carrier. In another embodiment, the invention features a pharmaceutical composition comprising chemically modified siRNA molecules of the invention targeting one or more genes in a pharmaceutically acceptable carrier. In another embodiment, the invention features a method

for treating or preventing a disease or condition in a patient, comprising administering to the patient a pharmaceutical composition of the invention under conditions suitable for the treatment or prevention of the disease or condition in the patient, alone or in conjunction with one or more other therapeutic compounds. In yet another embodiment, the invention features a method for reducing or preventing tissue rejection in a patient comprising administering to the patient a pharmaceutical composition of the invention under conditions suitable for the reduction or prevention of tissue rejection in the patient.

In another embodiment, the invention features a method for validating a gene target, comprising: (a) synthesizing a chemically modified siRNA molecule of the invention wherein one of the siRNA strands includes a sequence complimentary to RNA of a target gene; (b) introducing the chemically modified siRNA molecule into a cell, tissue, or organism under conditions suitable for modulating expression of the target gene in the cell, tissue, or organism; and (c) determining the function of the gene by assaying for any phenotypic change in the cell, tissue, or organism.

In one embodiment, the invention features a kit containing a chemically modified siRNA molecule of the invention that can be used to modulate the expression of a target gene in a cell, tissue, or organism. In another embodiment, the invention features a kit containing more than one chemically modified siRNA molecule of the invention that can be used to modulate the expression of more than one target gene in a cell, tissue, or organism.

In one embodiment, the invention features a cell containing one or more chemically modified siRNA molecules of the invention. In another embodiment, the cell containing a chemically modified siRNA molecule of the invention is a mammalian cell. In yet another embodiment, the cell containing a chemically modified siRNA molecule of the invention is a human cell.

In one embodiment, the synthesis of a chemically modified siRNA molecule of the invention comprises: (a) synthesis of two complimentary strands of the siRNA molecule; (b) annealing the two complimentary strands together under conditions suitable to obtain a double stranded siRNA molecule. In another embodiment, synthesis of the two complimentary strands of the siRNA molecule is by solid phase oligonucleotide synthesis. In yet another embodiment, synthesis of the two complimentary strands of the siRNA molecule is by solid phase tandem oligonucleotide synthesis.

In one embodiment, the invention features a method for synthesizing a siRNA duplex molecule comprising: (a) synthesizing a first oligonucleotide sequence strand of the siRNA molecule, wherein the first oligonucleotide sequence strand comprises a cleavable linker molecule that can be used as a scaffold for the synthesis of the second oligonucleotide sequence strand of the siRNA; (b) synthesizing the second oligonucleotide sequence strand of siRNA on the scaffold of the first oligonucleotide sequence strand, wherein the second oligonucleotide sequence strand further comprises a chemical moiety than can be used to purify the siRNA duplex; (c) cleaving the linker molecule of (a) under conditions suitable for the two siRNA oligonucleotide strands to hybridize and form a stable duplex; and (d) purifying the siRNA duplex utilizing the chemical moiety of the second oligonucleotide sequence strand. In another embodiment, cleavage of the linker molecule in (c) above takes place during deprotection of the oligonucleotide, for example under hydrolysis conditions using an alkylamine base such as methylamine. In another embodiment, the method of synthesis comprises solid phase synthesis on a solid support such as controlled pore glass (CPG) or polystyrene, wherein the first sequence of (a) is synthesized on a cleavable linker, such as a succinyl linker, using the solid support as a scaffold. The cleavable linker in (a) used as a scaffold for synthesizing the second strand can comprise similar reactivity as the solid support derivatized linker, such that cleavage of the solid support derivatized linker and the cleavable linker of (a) takes place concomitantly. In another embodiment, the chemical moiety of (b) that can used to isolate the attached oligonucleotide sequence comprises a trityl group, for example a dimethoxytrityl group, which can be employed in a trityl-on synthesis strategy as described herein. In yet another embodiment, the chemical moiety, such as a dimethoxytrityl group, is removed during purification, for example using acidic conditions.

In a further embodiment, the method for siRNA synthesis is a solution phase synthesis or hybrid phase synthesis wherein both strands of the siRNA duplex are synthesized in tandem using a cleavable linker attached to the first sequence which acts a scaffold for synthesis of the second sequence. Cleavage of the linker under conditions suitable for hybridization of the separate siRNA sequence strands results in formation of the double stranded siRNA molecule.

In another embodiment, the invention features a method for synthesizing a siRNA duplex molecule comprising: (a) synthesizing one oligonucleotide sequence strand of the siRNA molecule, wherein the sequence comprises a cleavable linker molecule that can be used as a scaffold for the synthesis of another oligonucleotide sequence; (b) synthesizing a second oligonucleotide sequence having complementarity to the first sequence strand on the scaffold of (a), wherein the second sequence comprises the other strand of the double stranded siRNA

molecule and wherein the second sequence further comprises a chemical moiety than can be used to isolate the attached oligonucleotide sequence; (c) purifying the product of (b) utilizing the chemical moiety of the second oligonucleotide sequence strand under conditions suitable for isolating the full length sequence comprising both siRNA oligonucleotide strands connected by the cleavable linker; and (d) under conditions suitable for the two siRNA oligonucleotide strands to hybridize and form a stable duplex. In another embodiment, cleavage of the linker molecule in (c) above takes place during deprotection of the oligonucleotide, for example under hydrolysis conditions. In another embodiment, cleavage of the linker molecule in (c) above takes place after deprotection of the oligonucleotide. In another embodiment, the method of synthesis comprises solid phase synthesis on a solid support such as controlled pore glass (CPG) or polystyrene, wherein the first sequence of (a) is synthesized on a cleavable linker, such as a succinyl linker, using the solid support as a scaffold. The cleavable linker in (a) used as a scaffold for synthesizing the second strand can comprise similar reactivity or differing reactivity as the solid support derivatized linker, such that cleavage of the solid support derivatized linker and the cleavable linker of (a) takes place either concomitantly or sequentially. In another embodiment, the chemical moiety of (b) that can used to isolate the attached oligonucleotide sequence comprises a trityl group, for example a dimethoxytrityl group.

In another embodiment, the invention features a method for making a double stranded siRNA molecule in a single synthetic process, comprising: (a) synthesizing an oligonucleotide having a first and a second sequence, wherein the first sequence is complimentary to the second sequence, and the first oligonucleotide sequence is linked to the second sequence via a cleavable linker, and wherein a terminal 5'-protecting group, for example a 5'-O-dimethoxytrityl group (5'-O-DMT) remains on the oligonucleotide having the second sequence; (b) deprotecting the oligonucleotide whereby the deprotection results in the cleavage of the linker joining the two oligonucleotide sequences; and (c) purifying the product of (b) under conditions suitable for isolating the double stranded siRNA molecule, for example using a trityl-on synthesis strategy as described herein.

In one embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications, for example one or more chemical modifications having Formula I, II, III, IV, or V, that increases the nuclease resistance of the siRNA construct.

In another embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications described herein

that modulates the binding affinity between the sense and antisense strands of the siRNA construct.

In another embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications described herein that modulates the binding affinity between the antisense strand of the siRNA construct and a complimentary target RNA sequence within a cell.

In one embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications described herein that modulates the polymerase activity of a cellular polymerase capable of generating additional endogenous siRNA molecules having sequence homology to the chemically modified siRNA construct.

In one embodiment, the invention features chemically modified siRNA constructs that mediate RNAi in a cell, wherein the chemical modifications do not significantly effect the interaction of siRNA with a target RNA molecule and/or proteins or other factors that are essential for RNAi in a manner that would decrease the efficacy of RNAi mediated by such siRNA constructs.

In another embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications described herein that modulates the cellular uptake of the siRNA construct.

In another embodiment, the invention features siRNA constructs that mediate RNAi, wherein the siRNA construct comprises one or more chemical modifications described herein that increases the bioavailability of the siRNA construct, for example by attaching polymeric conjugates such as polyethyleneglycol or equivalent conjugates that improve the pharmacokinetic of the siRNA construct, or by attaching conjugates that target specific tissue types or cell types *in vivo*. Non-limiting examples of such conjugates are described in Vargeese *et al.*, US Serial No. 60/311,865 incorporated by reference herein

In another embodiment, polyethylene glycol (PEG) can be covalently attached to siRNA compounds of the present invention. The attached PEG can be any molecular weight, preferably from about 2,000 to about 50,000 daltons (Da).

The term "short interfering RNA" or "siRNA" as used herein refers to a double stranded nucleic acid molecule capable of RNA interference "RNAi", see for example Bass, 2001,

Nature, 411, 428-429; Elbashir et al., 2001, *Nature*, 411, 494-498; and Kreutzer *et al.*, International PCT Publication No. WO 00/44895; Zernicka-Goetz *et al.*, International PCT Publication No. WO 01/36646; Fire, International PCT Publication No. WO 99/32619; Plaetinck *et al.*, International PCT Publication No. WO 00/01846; Mello and Fire, International PCT Publication No. WO 01/29058; Deschamps-Depaillette, International PCT Publication No. WO 99/07409; and Li *et al.*, International PCT Publication No. WO 00/44914. As used herein, siRNA molecules need not be limited to those molecules containing only RNA, but further encompasses chemically modified nucleotides and non-nucleotides.

By "modulate" is meant that the expression of the gene, or level of RNA molecule or equivalent RNA molecules encoding one or more proteins or protein subunits, or activity of one or more proteins or protein subunits is up regulated or down regulated, such that expression, level, or activity is greater than or less than that observed in the absence of the modulator. For example, the term "modulate" can mean "inhibit," but the use of the word "modulate" is not limited to this definition.

By "inhibit" it is meant that the activity of a gene expression product or level of RNAs or equivalent RNAs encoding one or more gene products is reduced below that observed in the absence of the nucleic acid molecule of the invention. In one embodiment, inhibition with a siRNA molecule preferably is below that level observed in the presence of an inactive or attenuated molecule that is unable to mediate an RNAi response. In another embodiment, inhibition of gene expression with the siRNA molecule of the instant invention is greater in the presence of the siRNA molecule than in its absence.

By "gene" or "target gene" is meant, a nucleic acid that encodes an RNA, for example, nucleic acid sequences including, but not limited to, structural genes encoding a polypeptide. The target gene can be a gene derived from a cell, an endogenous gene, a transgene, or exogenous genes such as genes of a pathogen, for example a virus, which is present in the cell after infection thereof. The cell containing the target gene can be derived from or contained in any organism, for example a plant, animal, protozoan, virus, bacterium, or fungus. Non-limiting examples of plants include monocots, dicots, or gymnosperms. Non-limiting examples of animals include vertebrates or invertebrates. Non-limiting examples of fungi include molds or yeasts.

By "complementarity" is meant that a nucleic acid can form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. In reference to the nucleic molecules of the present invention, the binding free energy for a nucleic

acid molecule with its complementary sequence is sufficient to allow the relevant function of the nucleic acid to proceed, e.g., RNAi activity. Determination of binding free energies for nucleic acid molecules is well known in the art (see, e.g., Turner et al., 1987, *CSH Symp. Quant. Biol.* LII pp.123-133; Frier et al., 1986, *Proc. Nat. Acad. Sci.* USA 83:9373-9377; Turner et al., 1987, *J. Am. Chem. Soc.* 109:3783-3785). A percent complementarity indicates the percentage of contiguous residues in a nucleic acid molecule that can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). "Perfectly complementary" means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence.

The siRNA molecules of the invention represent a novel therapeutic approach to treat a variety of pathologic indications, including cancer, infectious disease, cardiovascular, neurologic, inflammatory, immunologic, metabolic, endocrine, or genetic diseases and disorders or any other disease or condition that is based upon gene expression in humans, animals, plants, bacteria, and/or fungi.

In one embodiment of the present invention, each sequence of a siRNA molecule of the invention is independently 18 to 24 nucleotides in length, in specific embodiments about 18, 19, 20, 21, 22, 23, or 24 nucleotides in length. In another embodiment, the siRNA duplexes of the invention independently comprise between 17 and 23 base pairs. In yet another embodiment, siRNA molecules of the invention comprising hairpin or circular structures are 35 to 55 nucleotides in lengh, or 38-44 nucleotides in length and comprising 16-22 base pairs. Exemplary synthetic siRNA molecules of the invention are shown in Table I.

As used herein "cell" is used in its usual biological sense, and does not refer to an entire multicellular organism, e.g., specifically does not refer to a human. The cell can be present in an organism, e.g., birds, plants and mammals such as humans, cows, sheep, apes, monkeys, swine, dogs, and cats. The cell can be prokaryotic (e.g., bacterial cell) or eukaryotic (e.g., mammalian or plant cell). The cell can be of somatic or germ line origin, totipotent or pluripotent, dividing or non-dividing. The cell can also be derived from or can comprise a gamete or embryo, a stem cell, or a fully differentiated cell.

The siRNA molecules of the invention are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells or tissues. The nucleic acid or nucleic acid complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, infusion pump or stent, with or without their incorporation in

biopolymers. In particular embodiments, the nucleic acid molecules of the invention comprise sequences shown in Table I. Examples of such nucleic acid molecules consist essentially of sequences defined in this table.

In another aspect, the invention provides mammalian cells containing one or more siRNA molecules of this invention. The one or more siRNA molecules can independently be targeted to the same or different sites.

By "RNA" is meant a molecule comprising at least one ribonucleotide residue. By "ribonucleotide" is meant a nucleotide with a hydroxyl group at the 2' position of a β-D-ribofuranose moiety.

By "patient" is meant an organism, which is a donor or recipient of explanted cells or the cells themselves. "Patient" also refers to an organism to which the nucleic acid molecules of the invention can be administered. In one embodiment, a patient is a mammal or mammalian cells. In another embodiment, a patient is a human or human cells.

The term "phosphorothioate" as used herein refers to an internucleotide linkage having Formula I, wherein Z and/or W comprise a sulfur atom. Hence, the term phosphorothioate refers to both phosphorothioate and phosphorodithioate internucleotide linkages.

The term "universal base" as used herein refers to nucleotide base analogs that form base pairs with each of the natural DNA/RNA bases with little discrimination between them. Non-limiting examples of universal bases include C-phenyl, C-naphthyl and other aromatic derivatives, inosine, azole carboxamides, and nitroazole derivatives such as 3-nitropyrrole, 4-nitroindole, 5-nitroindole, and 6-nitroindole as known in the art (see for example Loakes, 2001, *Nucleic Acids Research*, 29, 2437-2447).

The nucleic acid molecules of the instant invention, individually, or in combination or in conjunction with other drugs, can be used to treat diseases or conditions discussed herein. For example, to treat a particular disease or condition, the siRNA molecules can be administered to a patient or can be administered to other appropriate cells evident to those skilled in the art, individually or in combination with one or more drugs under conditions suitable for the treatment.

In a further embodiment, the siRNA molecules can be used in combination with other known treatments to treat conditions or diseases discussed above. For example, the described molecules could be used in combination with one or more known therapeutic agents to treat a

disease or condition. Non-limiting examples of other therapeutic agents that can be readily combined with siRNA molecule of the invention are enzymatic nucleic acid molecules, allosteric nucleic acid molecules, antisense, decoy, or aptamer nucleic acid molecules, antibodies such as monoclonal antibodies, small molecules, and other organic and/or inorganic compounds including metals, salts and ions.

By "comprising" is meant including, but not limited to, whatever follows the word "comprising". Thus, use of the term "comprising" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of". Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

First the drawings will be described briefly.

Drawings

Figure 1 shows a non-limiting example of a scheme for the synthesis of siRNA molecules. The complimentary siRNA sequence strands, strand 1 and strand 2, are synthesized in tandem and are connected by a cleavable linkage, such as a nucleotide succinate or abasic succinate, which can be the same or different from the cleavable linker used for solid phase synthesis on a solid support. The synthesis can be either solid phase or solution phase, in the example shown, the synthesis is a solid phase synthesis. The synthesis is performed such that a protecting group, such as a dimethoxytrityl group, remains intact on the terminal nucleotide of the tandem oligonucleotide. Upon cleavage and deprotection of the oligonucleotide, the two siRNA strands spontaneously hybridize to form a siRNA duplex, which allows the purification of the duplex by utilizing the properties of the terminal protecting group, for example by applying a trityl on

purification method wherein only duplexes/oligonucleotides with the terminal protecting group are isolated.

Figure 2 shows a MALDI-TOV mass spectrum of a purified siRNA duplex synthesized by a method of the invention. The two peaks shown correspond to the predicted mass of the separate siRNA sequence strands. This result demonstrates that the siRNA duplex generated from tandem synthesis can be purified as a single entity using a simple trityl-on purification methodology.

Figure 3 shows the results of a-stability assay used to determine the serum stability of chemically modified siRNA constructs.

Figure 4 shows the results of an RNAi activity screen of phosphorothioate modified siRNA constructs using a luciferase reporter system.

Figure 5 shows the results of an RNAi activity screen of phosphorothioate and universal base modified siRNA constructs using a luciferase reporter system.

Figure 6 shows the results of an RNAi activity screen of 2'-O-methyl modified siRNA constructs using a luciferase reporter system.

Figure 7 shows the results of an RNAi activity screen of 2'-O-methyl and 2'-deoxy-2'-fluoro modified siRNA constructs using a luciferase reporter system.

Figure 8 shows the results of an RNAi activity screen of a phosphorothioate modified siRNA construct using a luciferase reporter system.

Figure 9 shows the results of an RNAi activity screen of an inverted deoxyabasic modified siRNA construct generated via tandem synthesis using a luciferase reporter system.

Figure 10 shows the results of a siRNA titration study wherein the RNAi activity of a phosphorothioate modified siRNA construct is compared to that of a siRNA construct consisting of all ribonucleotides except for two terminal thymidine residues.

Figure 11 shows a non-limiting example of siRNA constructs targeting viral replication of a HCV/poliovirus chimera.

Figure 12 shows a non-limiting example of a dose response study of a siRNA construct targeting viral replication of a HCV/poliovirus chimera.

Figure 13 shows a non-limiting example of a chemically modified siRNA construct targeting viral replication of a HCV/poliovirus chimera.

Figure 14 shows a non-limiting example of a chemically modified siRNA construct targeting viral replication of a HCV/poliovirus chimera.

Mechanism of action of Nucleic Acid Molecules of the Invention

RNA interference refers to the process of sequence specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA) (Fire *et al.*, 1998, *Nature*, 391, 806). The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla (Fire *et al.*, 1999, *Trends Genet.*, 15, 358). Such protection from foreign gene expression may have evolved in response to the production of double stranded RNAs (dsRNA) derived from viral infection or the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA) (Berstein *et al.*, 2001, *Nature*, 409, 363). Short interfering RNAs derived from dicer activity are typically about 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner *et al.*, 2001, *Science*, 293, 834). The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single stranded RNA having sequence homologous to the siRNA. Cleavage of the target RNA takes place in the middle of the region complementary to the guide sequence of the siRNA duplex (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188).

Short interfering RNA mediated RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. Elegans*. Wianny and Goetz, 1999,

Nature Cell Biol., 2, 70, describes RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in Drosophila embryonic lysates has revealed certain requirements for siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, substitution of one or both siRNA strands with 2'-deoxy or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of 3'-terminal siRNA nucleotides with deoxy nucleotides was shown to be tolerated. Mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA (Nykanen *et al.*, 2001, *Cell*, 107, 309), however siRNA molecules lacking a 5'-phosphate are active when introduced exogenously, suggesting that 5'-phosphorylation of siRNA constructs may occur *in vivo*.

Synthesis of Nucleic acid Molecules

Synthesis of nucleic acids greater than 100 nucleotides in length is difficult using automated methods, and the therapeutic cost of such molecules is prohibitive. In this invention, small nucleic acid motifs ("small" refers to nucleic acid motifs no more than 100 nucleotides in length, preferably no more than 80 nucleotides in length, and most preferably no more than 50 nucleotides in length; *e.g.*, individual siRNA oligonucleotide sequences or siRNA sequences synthesized in tandem, are preferably used for exogenous delivery. The simple structure of these molecules increases the ability of the nucleic acid to invade targeted regions of protein and/or RNA structure. Exemplary molecules of the instant invention are chemically synthesized, and others can similarly be synthesized.

Oligonucleotides (e.g., certain modified oligonucleotides or portions of oligonucleotides lacking ribonucleotides) are synthesized using protocols known in the art, for example as described in Caruthers *et al.*, 1992, *Methods in Enzymology* 211, 3-19, Thompson *et al.*, International PCT Publication No. WO 99/54459, Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684, Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, Brennan *et al.*, 1998, *Biotechnol Bioeng.*, 61, 33-45, and Brennan, US patent No. 6,001,311. All of these references are

incorporated herein by reference. The synthesis of oligonucleotides makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2 µmol scale protocol with a 2.5 min coupling step for 2'-O-methylated nucleotides and a 45 sec coupling step for 2'-deoxy nucleotides or 2'-deoxy-2'-fluoro nucleotides. Table II outlines the amounts and the contact times of the reagents used in the synthesis cycle. Alternatively, syntheses at the 0.2 µmol scale can be performed on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60 µL of 0.11 M = 6.6 µmol) of 2'-O-methyl phosphoramidite and a 105-fold excess of S-ethyl tetrazole (60 µL of 0.25 M = 15 µmol) can be used in each coupling cycle of 2'-O-methyl residues relative to polymer-bound 5'-hydroxyl. A 22-fold excess (40 µL of 0.11 M = 4.4 µmol) of deoxy phosphoramidite and a 70-fold excess of S-ethyl tetrazole (40 µL of 0.25 M = 10 µmol) can be used in each coupling cycle of deoxy residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); and oxidation solution is 16.9 mM I_2, 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide, 0.05 M in acetonitrile) is used.

Deprotection of the DNA-based oligonucleotides is performed as follows: the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to –20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H2O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder.

The method of synthesis used for normal RNA including certain siRNA molecules of the invention follows the procedure as described in Usman *et al.*, 1987, *J. Am. Chem. Soc.*, 109,

7845; Scaringe *et al.*, 1990, *Nucleic Acids Res.*, 18, 5433; and Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684 Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2 µmol scale protocol with a 7.5 min coupling step for alkylsilyl protected nucleotides and a 2.5 min coupling step for 2'-O-methylated nucleotides. Table II outlines the amounts and the contact times of the reagents used in the synthesis cycle. Alternatively, syntheses at the 0.2 µmol scale can be done on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60 µL of 0.11 M = 6.6 µmol) of 2'-O-methyl phosphoramidite and a 75-fold excess of S-ethyl tetrazole (60 µL of 0.25 M = 15 µmol) can be used in each coupling cycle of 2'-O-methyl residues relative to polymer-bound 5'-hydroxyl. A 66-fold excess (120 µL of 0.11 M = 13.2 µmol) of alkylsilyl (ribo) protected phosphoramidite and a 150-fold excess of S-ethyl tetrazole (120 µL of 0.25 M = 30 µmol) can be used in each coupling cycle of ribo residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); oxidation solution is 16.9 mM I_2, 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide0.05 M in acetonitrile) is used.

Deprotection of the RNA is performed using either a two-pot or one-pot protocol. For the two-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to –20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H2O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder. The base deprotected oligoribonucleotide is resuspended in anhydrous TEA/HF/NMP solution (300 µL of a solution of 1.5 mL N-methylpyrrolidinone, 750 µL TEA and 1 mL TEA•3HF to provide a 1.4 M HF concentration) and heated to 65 °C. After 1.5 h, the oligomer is quenched with 1.5 M NH_4HCO_3.

Alternatively, for the one-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 33% ethanolic methylamine/DMSO: 1/1 (0.8 mL) at 65 °C for 15 min. The vial is brought to r.t. TEA•3HF (0.1 mL) is added and the vial is heated at 65 °C for 15 min. The sample is cooled at –20 °C and then quenched with 1.5 M NH_4HCO_3.

For purification of the trityl-on oligomers, the quenched NH_4HCO_3 solution is loaded onto a C-18 containing cartridge that had been prewashed with acetonitrile followed by 50 mM TEAA. After washing the loaded cartridge with water, the RNA is detritylated with 0.5% TFA for 13 min. The cartridge is then washed again with water, salt exchanged with 1 M NaCl and washed with water again. The oligonucleotide is then eluted with 30% acetonitrile.

The average stepwise coupling yields are typically >98% (Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677-2684). Those of ordinary skill in the art will recognize that the scale of synthesis can be adapted to be larger or smaller than the example described above including but not limited to 96-well format, all that is important is the ratio of chemicals used in the reaction.

Alternatively, the nucleic acid molecules of the present invention can be synthesized separately and joined together post-synthetically, for example, by ligation (Moore *et al.*, 1992, *Science* 256, 9923; Draper *et al.*, International PCT publication No. WO 93/23569; Shabarova *et al.*, 1991, *Nucleic Acids Research* 19, 4247; Bellon *et al.*, 1997, *Nucleosides & Nucleotides*, 16, 951; Bellon *et al.*, 1997, *Bioconjugate Chem.* 8, 204), or by hybridization following synthesis and/or deprotection.

The siRNA molecules of the invention can also be synthesized via a tandem synthesis methodology as described in Example 1 herein, wherein both siRNA strands are synthesized as a contiguous oligonucleotide sequence separated by a cleavable linker which is subsequently cleaved to provide separate siRNA sequences that hybridize and permit purification of the siRNA duplex. The tandem synthesis of siRNA as described herein can be readily adapted to both multiwell/multiplate synthesis platforms such as 96 well or similarly larger multi-well platforms. The tandem synthesis of siRNA as described herein can also be readily adapted to large scale synthesis platforms employing batch reactors, synthesis columns and the like.

The nucleic acid molecules of the present invention can be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-*C*-allyl, 2'-fluoro, 2'-*O*-methyl, 2'-H (for a review see Usman and Cedergren, 1992, *TIBS* 17, 34; Usman *et al.*, 1994, *Nucleic Acids Symp. Ser.* 31, 163). siRNA constructs can be purified by gel

electrophoresis using general methods or can be purified by high pressure liquid chromatography (HPLC; see Wincott *et al., supra*, the totality of which is hereby incorporated herein by reference) and re-suspended in water.

The sequences of the siRNA constructs that are chemically synthesized, useful in this study, are shown in Table I. The siRNA construct sequences listed in Table I can be formed of ribonucleotides or other nucleotides or non-nucleotides

Optimizing Activity of the nucleic acid molecule of the invention.

Chemically synthesizing nucleic acid molecules with modifications (base, sugar and/or phosphate) can prevent their degradation by serum ribonucleases, which can increase their potency (see *e.g.*, Eckstein *et al.*, International Publication No. WO 92/07065; Perrault *et al.*, 1990 *Nature* 344, 565; Pieken et al., 1991, *Science* 253, 314; Usman and Cedergren, 1992, *Trends in Biochem. Sci.* 17, 334; Usman *et al.*, International Publication No. WO 93/15187; and Rossi *et al.*, International Publication No. WO 91/03162; Sproat, US Patent No. 5,334,711; Gold *et al.*, US 6,300,074; and Burgin *et al., supra;* all of which are incorporated by reference herein). All of the above references describe various chemical modifications that can be made to the base, phosphate and/or sugar moieties of the nucleic acid molecules described herein. Modifications that enhance their efficacy in cells, and removal of bases from nucleic acid molecules to shorten oligonucleotide synthesis times and reduce chemical requirements are desired.

There are several examples in the art describing sugar, base and phosphate modifications that can be introduced into nucleic acid molecules with significant enhancement in their nuclease stability and efficacy. For example, oligonucleotides are modified to enhance stability and/or enhance biological activity by modification with nuclease resistant groups, for example, 2'-amino, 2'-*C*-allyl, 2'-fluoro, 2'-*O*-methyl, 2'-O-allyl, 2'-H, nucleotide base modifications (for a review see Usman and Cedergren, 1992, *TIBS.* 17, 34; Usman *et al.*, 1994, *Nucleic Acids Symp. Ser.* 31, 163; Burgin *et al.*, 1996, *Biochemistry*, 35, 14090). Sugar modification of nucleic acid molecules have been extensively described in the art (see Eckstein *et al.*, *International Publication* PCT No. WO 92/07065; Perrault *et al. Nature*, 1990, *344*, 565-568; Pieken *et al. Science*, 1991, *253*, 314-317; Usman and Cedergren, *Trends in Biochem. Sci.* , 1992, *17*, 334-339; Usman *et al. International Publication* PCT No. WO 93/15187; Sproat, *US Patent* No. 5,334,711 and Beigelman *et al.*, 1995, *J. Biol. Chem.*, 270, 25702; Beigelman *et al.*, International PCT publication No. WO 97/26270; Beigelman *et al.*, US Patent No. 5,716,824; Usman *et al.*, US patent No. 5,627,053; Woolf *et al.*, International PCT Publication No. WO 98/13526; Thompson *et al.*, USSN 60/082,404 which was filed on April 20, 1998; Karpeisky *et al.*, 1998,

Tetrahedron Lett., 39, 1131; Earnshaw and Gait, 1998, *Biopolymers (Nucleic Acid Sciences)*, 48, 39-55; Verma and Eckstein, 1998, *Annu. Rev. Biochem.*, 67, 99-134; and Burlina *et al.*, 1997, *Bioorg. Med. Chem.*, 5, 1999-2010; all of the references are hereby incorporated in their totality by reference herein). Such publications describe general methods and strategies to determine the location of incorporation of sugar, base and/or phosphate modifications and the like into nucleic acid molecules without modulating catalysis, and are incorporated by reference herein. In view of such teachings, similar modifications can be used as described herein to modify the siRNA nucleic acid molecules of the instant invention so long as the ability of siRNA to promote RNAi is cells is not significantly inhibited..

While chemical modification of oligonucleotide internucleotide linkages with phosphorothioate, phosphorothioate, and/or 5'-methylphosphonate linkages improves stability, excessive modifications can cause some toxicity or decreased activity. Therefore, when designing nucleic acid molecules, the amount of these internucleotide linkages should be minimized. The reduction in the concentration of these linkages should lower toxicity, resulting in increased efficacy and higher specificity of these molecules.

Small interfering RNA (siRNA) molecules having chemical modifications that maintain or enhance activity are provided. Such a nucleic acid is also generally more resistant to nucleases than an unmodified nucleic acid. Accordingly, the *in vitro* and/or *in vivo* activity should not be significantly lowered. In cases in which modulation is the goal, therapeutic nucleic acid molecules delivered exogenously should optimally be stable within cells until translation of the target RNA has been modulated long enough to reduce the levels of the undesirable protein. This period of time varies between hours to days depending upon the disease state. Improvements in the chemical synthesis of RNA and DNA (Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677; Caruthers *et al.*, 1992, *Methods in Enzymology* 211,3-19 (incorporated by reference herein)) have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability, as described above.

In one embodiment, nucleic acid molecules of the invention include one or more G-clamp nucleotides. A G-clamp nucleotide is a modified cytosine analog wherein the modifications confer the ability to hydrogen bond both Watson-Crick and Hoogsteen faces of a complementary guanine within a duplex, see for example Lin and Matteucci, 1998, *J. Am. Chem. Soc.*, 120, 8531-8532. A single G-clamp analog substitution within an oligonucleotide can result in substantially enhanced helical thermal stability and mismatch discrimination when hybridized to

complementary oligonucleotides. The inclusion of such nucleotides in nucleic acid molecules of the invention results in both enhanced affinity and specificity to nucleic acid targets, complimentary seqnences, or template strands. In another embodiment, nucleic acid molecules of the invention include one or more LNA "locked nucleic acid" nucleotides such as a 2', 4'-C mythylene bicyclo nucleotide (see for example Wengel *et al.*, International PCT Publication No. WO 00/66604 and WO 99/14226).

In another embodiment, the invention features conjugates and/or complexes of siRNA molecules of the invention. Such conjugates and/or complexes can be used to facilitate delivery of siRNA molecules into a biological system, such as a cell. The conjugates and complexes provided by the instant invention can impart therapeutic activity by transferring therapeutic compounds across cellular membranes, altering the pharmacokinetics, and/or modulating the localization of nucleic acid molecules of the invention. The present invention encompasses the design and synthesis of novel conjugates and complexes for the delivery of molecules, including, but not limited to, small molecules, lipids, phospholipids, nucleosides, nucleotides, nucleic acids, antibodies, toxins, negatively charged polymers and other polymers, for example proteins, peptides, hormones, carbohydrates, polyethylene glycols, or polyamines, across cellular membranes. In general, the transporters described are designed to be used either individually or as part of a multi-component system, with or without degradable linkers. These compounds are expected to improve delivery and/or localization of nucleic acid molecules of the invention into a number of cell types originating from different tissues, in the presence or absence of serum (see Sullenger and Cech, US 5,854,038). Conjugates of the molecules described herein can be attached to biologically active molecules via linkers that are biodegradable, such as biodegradable nucleic acid linker molecules.

The term "biodegradable nucleic acid linker molecule" as used herein, refers to a nucleic acid molecule that is designed as a biodegradable linker to connect one molecule to another molecule, for example, a biologically active molecule. The stability of the biodegradable nucleic acid linker molecule can be modulated by using various combinations of ribonucleotides, deoxyribonucleotides, and chemically modified nucleotides, for example, 2'-O-methyl, 2'-fluoro, 2'-amino, 2'-O-amino, 2'-C-allyl, 2'-O-allyl, and other 2'-modified or base modified nucleotides. The biodegradable nucleic acid linker molecule can be a dimer, trimer, tetramer or longer nucleic acid molecule, for example, an oligonucleotide of about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length, or can comprise a single nucleotide with a phosphorus-based linkage, for example, a phosphoramidate or phosphodiester linkage.

The biodegradable nucleic acid linker molecule can also comprise nucleic acid backbone, nucleic acid sugar, or nucleic acid base modifications.

The term "biodegradable" as used herein, refers to degradation in a biological system, for example enzymatic degradation or chemical degradation.

The term "biologically active molecule" as used herein, refers to compounds or molecules that are capable of eliciting or modifying a biological response in a system. Non-limiting examples of biologically active siRNA molecules either alone or in combination with othe molecules contemplated by the instant invention include therapeutically active molecules such as antibodies, hormones, antivirals, peptides, proteins, chemotherapeutics, small molecules, vitamins, co-factors, nucleosides, nucleotides, oligonucleotides, enzymatic nucleic acids, antisense nucleic acids, triplex forming oligonucleotides, 2,5-A chimeras, siRNA, dsRNA, allozymes, aptamers, decoys and analogs thereof. Biologically active molecules of the invention also include molecules capable of modulating the pharmacokinetics and/or pharmacodynamics of other biologically active molecules, for example, lipids and polymers such as polyamines, polyamides, polyethylene glycol and other polyethers.

The term "phospholipid" as used herein, refers to a hydrophobic molecule comprising at least one phosphorus group. For example, a phospholipid can comprise a phosphorus-containing group and saturated or unsaturated alkyl group, optionally substituted with OH, COOH, oxo, amine, or substituted or unsubstituted aryl groups.

Therapeutic nucleic acid molecules (*e.g.*, siRNA molecules) delivered exogenously optimally are stable within cells until reverse trascription of the RNA has been modulated long enough to reduce the levels of the RNA transcript. The nucleic acid molecules are resistant to nucleases in order to function as effective intracellular therapeutic agents. Improvements in the chemical synthesis of nucleic acid molecules described in the instant invention and in the art have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability as described above.

In yet another embodiment, siRNA molecules having chemical modifications that maintain or enhance enzymatic activity of proteins involved in RNAi are provided. Such nucleic acids are also generally more resistant to nucleases than unmodified nucleic acids. Thus, *in vitro* and/or *in vivo* the activity should not be significantly lowered.

Use of the nucleic acid-based molecules of the invention will lead to better treatment of the disease progression by affording the possibility of combination therapies (*e.g.*, multiple siRNA molecules targeted to different genes; nucleic acid molecules coupled with known small molecule modulators; or intermittent treatment with combinations of molecules, including different motifs and/or other chemical or biological molecules). The treatment of patients with siRNA molecules can also include combinations of different types of nucleic acid molecules, such as enzymatic nucleic acid molecules (ribozymes), allozymes, antisense, 2,5-A oligoadenylate, decoys, aptamers etc.

In another aspect a siRNA molecule of the invention comprises one or more 5' and/or a 3'- cap structure, for example on only the sense siRNA strand, antisense siRNA strand, or both siRNA strands.

By "cap structure" is meant chemical modifications, which have been incorporated at either terminus of the oligonucleotide (see, for example, Adamic *et al.*, US 5,998,203, incorporated by reference herein). These terminal modifications protect the nucleic acid molecule from exonuclease degradation, and may help in delivery and/or localization within a cell. The cap may be present at the 5'-terminus (5'-cap) or at the 3'-terminal (3'-cap) or may be present on both termini. In non-limiting examples: the 5'-cap is selected from the group comprising inverted abasic residue (moiety); 4',5'-methylene nucleotide; 1-(beta-D-erythrofuranosyl) nucleotide, 4'-thio nucleotide; carbocyclic nucleotide; 1,5-anhydrohexitol nucleotide; L-nucleotides; alpha-nucleotides; modified base nucleotide; phosphorodithioate linkage; *threo*-pentofuranosyl nucleotide; acyclic 3',4'-seco nucleotide; acyclic 3,4-dihydroxybutyl nucleotide; acyclic 3,5-dihydroxypentyl nucleotide, 3'-3'-inverted nucleotide moiety; 3'-3'-inverted abasic moiety; 3'-2'-inverted nucleotide moiety; 3'-2'-inverted abasic moiety; 1,4-butanediol phosphate; 3'-phosphoramidate; hexylphosphate; aminohexyl phosphate; 3'-phosphate; 3'-phosphorothioate; phosphorodithioate; or bridging or non-bridging methylphosphonate moiety.

In yet another preferred embodiment, the 3'-cap is selected from a group comprising, 4',5'-methylene nucleotide; 1-(beta-D-erythrofuranosyl) nucleotide; 4'-thio nucleotide, carbocyclic nucleotide; 5'-amino-alkyl phosphate; 1,3-diamino-2-propyl phosphate; 3-aminopropyl phosphate; 6-aminohexyl phosphate; 1,2-aminododecyl phosphate; hydroxypropyl phosphate; 1,5-anhydrohexitol nucleotide; L-nucleotide; alpha-nucleotide; modified base nucleotide; phosphorodithioate; *threo*-pentofuranosyl nucleotide; acyclic 3',4'-seco nucleotide; 3,4-dihydroxybutyl nucleotide; 3,5-dihydroxypentyl nucleotide, 5'-5'-inverted nucleotide moiety; 5'-

5'-inverted abasic moiety; 5'-phosphoramidate; 5'-phosphorothioate; 1,4-butanediol phosphate; 5'-amino; bridging and/or non-bridging 5'-phosphoramidate, phosphorothioate and/or phosphorodithioate, bridging or non bridging methylphosphonate and 5'-mercapto moieties (for more details see Beaucage and Iyer, 1993, *Tetrahedron* 49, 1925; incorporated by reference herein).

By the term "non-nucleotide" is meant any group or compound which can be incorporated into a nucleic acid chain in the place of one or more nucleotide units, including either sugar and/or phosphate substitutions, and allows the remaining bases to exhibit their enzymatic activity. The group or compound is abasic in that it does not contain a commonly recognized nucleotide base, such as adenosine, guanine, cytosine, uracil or thymine and therefore lacks a base at the 1'-position.

An "alkyl" group refers to a saturated aliphatic hydrocarbon, including straight-chain, branched-chain, and cyclic alkyl groups. Preferably, the alkyl group has 1 to 12 carbons. More preferably, it is a lower alkyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkyl group can be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO_2 or $N(CH_3)_2$, amino, or SH. The term also includes alkenyl groups that are unsaturated hydrocarbon groups containing at least one carbon-carbon double bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkenyl group has 1 to 12 carbons. More preferably, it is a lower alkenyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkenyl group may be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO_2, halogen, $N(CH_3)_2$, amino, or SH. The term "alkyl" also includes alkynyl groups that have an unsaturated hydrocarbon group containing at least one carbon-carbon triple bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkynyl group has 1 to 12 carbons. More preferably, it is a lower alkynyl of from 1 to 7 carbons, more preferably 1 to 4 carbons. The alkynyl group may be substituted or unsubstituted. When substituted the substituted group(s) is preferably, hydroxyl, cyano, alkoxy, =O, =S, NO_2 or $N(CH_3)_2$, amino or SH.

Such alkyl groups may also include aryl, alkylaryl, carbocyclic aryl, heterocyclic aryl, amide and ester groups. An "aryl" group refers to an aromatic group that has at least one ring having a conjugated pi electron system and includes carbocyclic aryl, heterocyclic aryl and biaryl groups, all of which may be optionally substituted. The preferred substituent(s) of aryl groups are halogen, trihalomethyl, hydroxyl, SH, OH, cyano, alkoxy, alkyl, alkenyl, alkynyl, and amino

groups. An "alkylaryl" group refers to an alkyl group (as described above) covalently joined to an aryl group (as described above). Carbocyclic aryl groups are groups wherein the ring atoms on the aromatic ring are all carbon atoms. The carbon atoms are optionally substituted. Heterocyclic aryl groups are groups having from 1 to 3 heteroatoms as ring atoms in the aromatic ring and the remainder of the ring atoms are carbon atoms. Suitable heteroatoms include oxygen, sulfur, and nitrogen, and include furanyl, thienyl, pyridyl, pyrrolyl, N-lower alkyl pyrrolo, pyrimidyl, pyrazinyl, imidazolyl and the like, all optionally substituted. An "amide" refers to an -C(O)-NH-R, where R is either alkyl, aryl, alkylaryl or hydrogen. An "ester" refers to an -C(O)-OR', where R is either alkyl, aryl, alkylaryl or hydrogen.

By "nucleotide" as used herein is as recognized in the art to include natural bases (standard), and modified bases well known in the art. Such bases are generally located at the 1' position of a nucleotide sugar moiety. Nucleotides generally comprise a base, sugar and a phosphate group. The nucleotides can be unmodified or modified at the sugar, phosphate and/or base moiety, (also referred to interchangeably as nucleotide analogs, modified nucleotides, non-natural nucleotides, non-standard nucleotides and other; see, for example, Usman and McSwiggen, *supra*; Eckstein *et al.*, International PCT Publication No. WO 92/07065; Usman *et al.*, International PCT Publication No. WO 93/15187; Uhlman & Peyman, *supra*, all are hereby incorporated by reference herein). There are several examples of modified nucleic acid bases known in the art as summarized by Limbach *et al.*, 1994, *Nucleic Acids Res.* 22, 2183. Some of the non-limiting examples of base modifications that can be introduced into nucleic acid molecules include, inosine, purine, pyridin-4-one, pyridin-2-one, phenyl, pseudouracil, 2, 4, 6-trimethoxy benzene, 3-methyl uracil, dihydrouridine, naphthyl, aminophenyl, 5-alkylcytidines (*e.g.*, 5-methylcytidine), 5-alkyluridines (*e.g.*, ribothymidine), 5-halouridine (*e.g.*, 5-bromouridine) or 6-azapyrimidines or 6-alkylpyrimidines (*e.g.* 6-methyluridine), propyne, and others (Burgin *et al.*, 1996, *Biochemistry*, 35, 14090; Uhlman & Peyman, *supra*). By "modified bases" in this aspect is meant nucleotide bases other than adenine, guanine, cytosine and uracil at 1' position or their equivalents.

In one embodiment, the invention features modified siRNA molecules, with phosphate backbone modifications comprising one or more phosphorothioate, phosphorodithioate, methylphosphonate, morpholino, amidate carbamate, carboxymethyl, acetamidate, polyamide, sulfonate, sulfonamide, sulfamate, formacetal, thioformacetal, and/or alkylsilyl, substitutions. For a review of oligonucleotide backbone modifications, see Hunziker and Leumann, 1995, *Nucleic Acid Analogues: Synthesis and Properties*, in *Modern Synthetic Methods*, VCH, 331-

417, and Mesmaeker *et al.*, 1994, *Novel Backbone Replacements for Oligonucleotides*, in *Carbohydrate Modifications in Antisense Research*, ACS, 24-39.

By "abasic" is meant sugar moieties lacking a base or having other chemical groups in place of a base at the 1' position, see for example Adamic *et al.*, US 5,998,203.

By "unmodified nucleoside" is meant one of the bases adenine, cytosine, guanine, thymine, uracil joined to the 1' carbon of β-D-ribo-furanose.

By "modified nucleoside" is meant any nucleotide base which contains a modification in the chemical structure of an unmodified nucleotide base, sugar and/or phosphate.

In connection with 2'-modified nucleotides as described for the present invention, by "amino" is meant 2'-NH_2 or 2'-*O*-NH_2, which may be modified or unmodified. Such modified groups are described, for example, in Eckstein *et al.*, U.S. 5,672,695 and Matulic-Adamic *et al.*, US 6,248,878, which are both incorporated by reference in their entireties.

Various modifications to nucleic acid siRNA structure can be made to enhance the utility of these molecules. Such modifications will enhance shelf-life, half-life *in vitro*, stability, and ease of introduction of such oligonucleotides to the target site, e.g., to enhance penetration of cellular membranes, and confer the ability to recognize and bind to targeted cells.

Administration of Nucleic Acid Molecules

Methods for the delivery of nucleic acid molecules are described in Akhtar *et al.*, 1992, *Trends Cell Bio.*, 2, 139; *Delivery Strategies for Antisense Oligonucleotide Therapeutics*, ed. Akhtar, 1995, Maurer *et al.*, 1999, *Mol. Membr. Biol.*, 16, 129-140; Hofland and Huang, 1999, *Handb. Exp. Pharmacol.*, 137, 165-192; and Lee *et al.*, 2000, *ACS Symp. Ser.*, 752, 184-192, all of which are incorporated herein by reference. Sullivan *et al.*, PCT WO 94/02595, further describes the general methods for delivery of nucleic acid molecules. These protocols can be utilized for the delivery of virtually any nucleic acid molecule. Nucleic acid molecules can be administered to cells by a variety of methods known to those of skill in the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres, or by proteinaceous vectors (O'Hare and Normand, International PCT Publication No. WO 00/53722). Alternatively, the nucleic acid/vehicle combination is locally delivered by direct injection or by use of an infusion pump. Direct injection of the nucleic acid molecules of the invention, whether subcutaneous, intramuscular, or intradermal, can take place using

standard needle and syringe methodologies, or by needle-free technologies such as those described in Conry *et al.*, 1999, *Clin. Cancer Res.*, 5, 2330-2337 and Barry *et al.*, International PCT Publication No. WO 99/31262. The molecules of the instant invention can be used as pharmaceutical agents. Pharmaceutical agents prevent, modulate the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state in a patient.

Thus, the invention features a pharmaceutical composition comprising one or more nucleic acid(s) of the invention in an acceptable carrier, such as a stabilizer, buffer, and the like. The polynucleotides of the invention can be administered (*e.g.*, RNA, DNA or protein) and introduced into a patient by any standard means, with or without stabilizers, buffers, and the like, to form a pharmaceutical composition. When it is desired to use a liposome delivery mechanism, standard protocols for formation of liposomes can be followed. The compositions of the present invention may also be formulated and used as tablets, capsules or elixirs for oral administration, suppositories for rectal administration, sterile solutions, suspensions for injectable administration, and the other compositions known in the art.

The present invention also includes pharmaceutically acceptable formulations of the compounds described. These formulations include salts of the above compounds, *e.g.*, acid addition salts, for example, salts of hydrochloric, hydrobromic, acetic acid, and benzene sulfonic acid.

A pharmacological composition or formulation refers to a composition or formulation in a form suitable for administration, *e.g.*, systemic administration, into a cell or patient, including for example a human. Suitable forms, in part, depend upon the use or the route of entry, for example oral, transdermal, or by injection. Such forms should not prevent the composition or formulation from reaching a target cell (*i.e.*, a cell to which the negatively charged nucleic acid is desirable for delivery). For example, pharmacological compositions injected into the blood stream should be soluble. Other factors are known in the art, and include considerations such as toxicity and forms that prevent the composition or formulation from exerting its effect.

By "systemic administration" is meant *in vivo* systemic absorption or accumulation of drugs in the blood stream followed by distribution throughout the entire body. Administration routes which lead to systemic absorption include, without limitation: intravenous, subcutaneous, intraperitoneal, inhalation, oral, intrapulmonary and intramuscular. Each of these administration routes exposes the siRNA molecules of the invention to an accessible diseased tissue. The rate of entry of a drug into the circulation has been shown to be a function of molecular weight or size.

The use of a liposome or other drug carrier comprising the compounds of the instant invention can potentially localize the drug, for example, in certain tissue types, such as the tissues of the reticular endothelial system (RES). A liposome formulation that can facilitate the association of drug with the surface of cells, such as, lymphocytes and macrophages is also useful. This approach may provide enhanced delivery of the drug to target cells by taking advantage of the specificity of macrophage and lymphocyte immune recognition of abnormal cells, such as cancer cells.

By "pharmaceutically acceptable formulation" is meant, a composition or formulation that allows for the effective distribution of the nucleic acid molecules of the instant invention in the physical location most suitable for their desired activity. Nonlimiting examples of agents suitable for formulation with the nucleic acid molecules of the instant invention include: P-glycoprotein inhibitors (such as Pluronic P85), which can enhance entry of drugs into the CNS (Jolliet-Riant and Tillement, 1999, *Fundam. Clin. Pharmacol.*, 13, 16-26); biodegradable polymers, such as poly (DL-lactide-coglycolide) microspheres for sustained release delivery after intracerebral implantation (Emerich, DF *et al*, 1999, *Cell Transplant*, 8, 47-58) (Alkermes, Inc. Cambridge, MA); and loaded nanoparticles, such as those made of polybutylcyanoacrylate, which can deliver drugs across the blood brain barrier and can alter neuronal uptake mechanisms (*Prog Neuropsychopharmacol Biol Psychiatry*, 23, 941-949, 1999). Other non-limiting examples of delivery strategies for the nucleic acid molecules of the instant invention include material described in Boado *et al.*, 1998, *J. Pharm. Sci.*, 87, 1308-1315; Tyler *et al.*, 1999, *FEBS Lett.*, 421, 280-284; Pardridge *et al.*, 1995, *PNAS USA.*, 92, 5592-5596; Boado, 1995, *Adv. Drug Delivery Rev.*, 15, 73-107; Aldrian-Herrada *et al.*, 1998, *Nucleic Acids Res.*, 26, 4910-4916; and Tyler *et al.*, 1999, *PNAS USA.*, 96, 7053-7058.

The invention also features the use of the composition comprising surface-modified liposomes containing poly (ethylene glycol) lipids (PEG-modified, or long-circulating liposomes or stealth liposomes). These formulations offer a method for increasing the accumulation of drugs in target tissues. This class of drug carriers resists opsonization and elimination by the mononuclear phagocytic system (MPS or RES), thereby enabling longer blood circulation times and enhanced tissue exposure for the encapsulated drug (Lasic *et al. Chem. Rev.* 1995, 95, 2601-2627; Ishiwata *et al.*, *Chem. Pharm. Bull.* 1995, 43, 1005-1011). Such liposomes have been shown to accumulate selectively in tumors, presumably by extravasation and capture in the neovascularized target tissues (Lasic *et al.*, *Science* 1995, 267, 1275-1276; Oku *et al.*,1995, *Biochim. Biophys. Acta*, 1238, 86-90). The long-circulating liposomes enhance the pharmacokinetics and pharmacodynamics of DNA and RNA, particularly compared to

conventional cationic liposomes which are known to accumulate in tissues of the MPS (Liu *et al.*, *J. Biol. Chem.* 1995, 42, 24864-24870; Choi *et al.*, International PCT Publication No. WO 96/10391; Ansell *et al.*, International PCT Publication No. WO 96/10390; Holland *et al.*, International PCT Publication No. WO 96/10392). Long-circulating liposomes are also likely to protect drugs from nuclease degradation to a greater extent compared to cationic liposomes, based on their ability to avoid accumulation in metabolically aggressive MPS tissues such as the liver and spleen.

The present invention also includes compositions prepared for storage or administration, which include a pharmaceutically effective amount of the desired compounds in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remington's Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro edit. 1985) hereby incorporated by reference herein. For example, preservatives, stabilizers, dyes and flavoring agents may be provided. These include sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid. In addition, antioxidants and suspending agents may be used.

The present invention also includes compositions prepared for storage or administration that include a pharmaceutically effective amount of the desired compounds in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remington's Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro edit. 1985), hereby incorporated by reference herein. For example, preservatives, stabilizers, dyes and flavoring agents can be provided. These include sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid. In addition, antioxidants and suspending agents can be used.

A pharmaceutically effective dose is that dose required to prevent, inhibit the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state. The pharmaceutically effective dose depends on the type of disease, the composition used, the route of administration, the type of mammal being treated, the physical characteristics of the specific mammal under consideration, concurrent medication, and other factors that those skilled in the medical arts will recognize. Generally, an amount between 0.1 mg/kg and 100 mg/kg body weight/day of active ingredients is administered dependent upon potency of the negatively charged polymer.

The nucleic acid molecules of the invention and formulations thereof can be administered orally, topically, parenterally, by inhalation or spray, or rectally in dosage unit formulations

containing conventional non-toxic pharmaceutically acceptable carriers, adjuvants and/or vehicles. The term parenteral as used herein includes percutaneous, subcutaneous, intravascular (*e.g.*, intravenous), intramuscular, or intrathecal injection or infusion techniques and the like. In addition, there is provided a pharmaceutical formulation comprising a nucleic acid molecule of the invention and a pharmaceutically acceptable carrier. One or more nucleic acid molecules of the invention can be present in association with one or more non-toxic pharmaceutically acceptable carriers and/or diluents and/or adjuvants, and if desired other active ingredients. The pharmaceutical compositions containing nucleic acid molecules of the invention can be in a form suitable for oral use, for example, as tablets, troches, lozenges, aqueous or oily suspensions, dispersible powders or granules, emulsion, hard or soft capsules, or syrups or elixirs.

Compositions intended for oral use can be prepared according to any method known to the art for the manufacture of pharmaceutical compositions and such compositions can contain one or more such sweetening agents, flavoring agents, coloring agents or preservative agents in order to provide pharmaceutically elegant and palatable preparations. Tablets contain the active ingredient in admixture with non-toxic pharmaceutically acceptable excipients that are suitable for the manufacture of tablets. These excipients can be, for example, inert diluents; such as calcium carbonate, sodium carbonate, lactose, calcium phosphate or sodium phosphate; granulating and disintegrating agents, for example, corn starch, or alginic acid; binding agents, for example starch, gelatin or acacia; and lubricating agents, for example magnesium stearate, stearic acid or talc. The tablets can be uncoated or they can be coated by known techniques. In some cases such coatings can be prepared by known techniques to delay disintegration and absorption in the gastrointestinal tract and thereby provide a sustained action over a longer period. For example, a time delay material such as glyceryl monostearate or glyceryl distearate can be employed.

Formulations for oral use can also be presented as hard gelatin capsules wherein the active ingredient is mixed with an inert solid diluent, for example, calcium carbonate, calcium phosphate or kaolin, or as soft gelatin capsules wherein the active ingredient is mixed with water or an oil medium, for example peanut oil, liquid paraffin or olive oil.

Aqueous suspensions contain the active materials in admixture with excipients suitable for the manufacture of aqueous suspensions. Such excipients are suspending agents, for example sodium carboxymethylcellulose, methylcellulose, hydropropyl-methylcellulose, sodium alginate, polyvinylpyrrolidone, gum tragacanth and gum acacia; dispersing or wetting agents can be a naturally-occurring phosphatide, for example, lecithin, or condensation products of an alkylene

oxide with fatty acids, for example polyoxyethylene stearate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate. The aqueous suspensions can also contain one or more preservatives, for example ethyl, or n-propyl p-hydroxybenzoate, one or more coloring agents, one or more flavoring agents, and one or more sweetening agents, such as sucrose or saccharin.

Oily suspensions can be formulated by suspending the active ingredients in a vegetable oil, for example arachis oil, olive oil, sesame oil or coconut oil, or in a mineral oil such as liquid paraffin. The oily suspensions can contain a thickening agent, for example beeswax, hard paraffin or cetyl alcohol. Sweetening agents and flavoring agents can be added to provide palatable oral preparations. These compositions can be preserved by the addition of an anti-oxidant such as ascorbic acid.

Dispersible powders and granules suitable for preparation of an aqueous suspension by the addition of water provide the active ingredient in admixture with a dispersing or wetting agent, suspending agent and one or more preservatives. Suitable dispersing or wetting agents or suspending agents are exemplified by those already mentioned above. Additional excipients, for example sweetening, flavoring and coloring agents, can also be present.

Pharmaceutical compositions of the invention can also be in the form of oil-in-water emulsions. The oily phase can be a vegetable oil or a mineral oil or mixtures of these. Suitable emulsifying agents can be naturally-occurring gums, for example gum acacia or gum tragacanth, naturally-occurring phosphatides, for example soy bean, lecithin, and esters or partial esters derived from fatty acids and hexitol, anhydrides, for example sorbitan monooleate, and condensation products of the said partial esters with ethylene oxide, for example polyoxyethylene sorbitan monooleate. The emulsions can also contain sweetening and flavoring agents.

Syrups and elixirs can be formulated with sweetening agents, for example glycerol, propylene glycol, sorbitol, glucose or sucrose. Such formulations can also contain a demulcent, a preservative and flavoring and coloring agents. The pharmaceutical compositions can be in the form of a sterile injectable aqueous or oleaginous suspension. This suspension can be formulated according to the known art using those suitable dispersing or wetting agents and suspending agents that have been mentioned above. The sterile injectable preparation can also

be a sterile injectable solution or suspension in a non-toxic parentally acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that can be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any bland fixed oil can be employed including synthetic mono-or diglycerides. In addition, fatty acids such as oleic acid find use in the preparation of injectables.

The nucleic acid molecules of the invention can also be administered in the form of suppositories, *e.g.*, for rectal administration of the drug. These compositions can be prepared by mixing the drug with a suitable non-irritating excipient that is solid at ordinary temperatures but liquid at the rectal temperature and will therefore melt in the rectum to release the drug. Such materials include cocoa butter and polyethylene glycols.

Nucleic acid molecules of the invention can be administered parenterally in a sterile medium. The drug, depending on the vehicle and concentration used, can either be suspended or dissolved in the vehicle. Advantageously, adjuvants such as local anesthetics, preservatives and buffering agents can be dissolved in the vehicle.

Dosage levels of the order of from about 0.1 mg to about 140 mg per kilogram of body weight per day are useful in the treatment of the above-indicated conditions (about 0.5 mg to about 7 g per patient per day). The amount of active ingredient that can be combined with the carrier materials to produce a single dosage form varies depending upon the host treated and the particular mode of administration. Dosage unit forms generally contain between from about 1 mg to about 500 mg of an active ingredient.

It is understood that the specific dose level for any particular patient depends upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, sex, diet, time of administration, route of administration, and rate of excretion, drug combination and the severity of the particular disease undergoing therapy.

For administration to non-human animals, the composition can also be added to the animal feed or drinking water. It can be convenient to formulate the animal feed and drinking water compositions so that the animal takes in a therapeutically appropriate quantity of the composition along with its diet. It can also be convenient to present the composition as a premix for addition to the feed or drinking water.

The nucleic acid molecules of the present invention may also be administered to a patient in combination with other therapeutic compounds to increase the overall therapeutic effect. The use of multiple compounds to treat an indication may increase the beneficial effects while reducing the presence of side effects.

In one embodiment, the invention compositions suitable for administering nucleic acid molecules of the invention to specific cell types, such as hepatocytes. For example, the asialoglycoprotein receptor (ASGPr) (Wu and Wu, 1987, *J. Biol. Chem.* 262, 4429-4432) is unique to hepatocytes and binds branched galactose-terminal glycoproteins, such as asialoorosomucoid (ASOR). Binding of such glycoproteins or synthetic glycoconjugates to the receptor takes place with an affinity that strongly depends on the degree of branching of the oligosaccharide chain, for example, triatennary structures are bound with greater affinity than biatennary or monoatennary chains (Baenziger and Fiete, 1980, *Cell*, 22, 611-620; Connolly *et al.*, 1982, *J. Biol. Chem.*, 257, 939-945). Lee and Lee, 1987, *Glycoconjugate J.*, 4, 317-328, obtained this high specificity through the use of N-acetyl-D-galactosamine as the carbohydrate moiety, which has higher affinity for the receptor, compared to galactose. This "clustering effect" has also been described for the binding and uptake of mannosyl-terminating glycoproteins or glycoconjugates (Ponpipom *et al.*, 1981, *J. Med. Chem.*, 24, 1388-1395). The use of galactose and galactosamine based conjugates to transport exogenous compounds across cell membranes can provide a targeted delivery approach to the treatment of liver disease such as HBV infection or hepatocellular carcinoma. The use of bioconjugates can also provide a reduction in the required dose of therapeutic compounds required for treatment. Furthermore, therapeutic bioavialability, pharmacodynamics, and pharmacokinetic parameters can be modulated through the use of nucleic acid bioconjugates of the invention.

Examples:

The following are non-limiting examples showing the selection, isolation, synthesis and activity of nucleic acids of the instant invention.

Example 1: Tandem synthesis of siRNA constructs

Exemplary siRNA molecules of the invention are synthesized in tandem using a cleavable linker, for example a succinyl-based linker. Tandem synthesis as described herein is followed by a one step purification process that provides RNAi molecules in high yield. This approach is highly amenable to siRNA synthesis in support of high throughput RNAi screening, and can be readily adapted to multi-column or multi-well synthesis platforms.

After completing a tandem synthesis of an siRNA oligo and its compliment in which the 5'-terminal dimethoxytrityl (5'-O-DMT) group remains intact (trityl on synthesis), the oligonucleotides are deprotected as described above. Following deprotection, the siRNA sequence strands are allowed to spontaneously hybridize. This hybridization yields a duplex in which one strand has retained the 5'-O-DMT group while the complimentary strand comprises a terminal 5'-hydroxyl. The newly formed duplex to behaves as a single molecule during routine solid-phase extraction purification (Trityl-On purification) even though only one molecule has a dimethoxytrityl group. Because the strands form a stable duplex, this dimethoxytrityl group (or an equivalent group, such as other trityl groups or other hydrophobic moieties) is all that is required to purify the pair of oligos, for example by using a C18 cartridge.

Standard phosphoramidite synthesis chemistry is used up to point of introducing a tandem linker, such as an inverted deoxyabasic succinate linker (see Figure 1) or an equivalent cleavable linker. A non-limiting example of linker coupling conditions that can be used includes a hindered base such as diisopropylethylamine (DIPA) and/or DMAP in the presence of an activator reagent such as Bromotripyrrolidinophosphoniumhexafluorophosphate (PyBrOP). After the linker is coupled, standard synthesis chemistry is utilized to complete synthesis of the second sequence leaving the terminal the 5'-O-DMT intact. Following synthesis, the resulting oligonucleotide is deprotected according to the procedures described herein and quenched with a suitable buffer, for example with 50mM NaOAc or 1.5M $NH_4H_2CO_3$.

Purification of the siRNA duplex can be readily accomplished using solid phase extraction, for example using a Waters C18 SepPak 1g cartridge conditioned with 1 column volume (CV) of acetonitrile, 2 CV H2O, and 2 CV 50mM NaOAc. The sample is loaded and then washed with 1 CV H2O or 50mM NaOAc. Failure sequences are eluted with 1 CV 14% ACN (Aqueous with 50mM NaOAc and 50mM NaCl). The column is then washed, for example with 1 CV H2O followed by on-column detritylation, for example by passing 1 CV of 1% aqueous trifluoroacetic acid (TFA) over the column, then adding a second CV of 1% aqueous TFA to the column and allowing to stand for approx. 10 minutes. The remaining TFA solution is removed and the column washed with H20 followed by 1 CV 1M NaCl and additional H2O. The siRNA duplex product is then eluted, for example using 1 CV 20% aqueous CAN.

Figure 2 provides an example of MALDI-TOF mass spectrometry analysis of a purified siRNA construct in which each peak corresponds to the calculated mass of an individual siRNA strand of the siRNA duplex. The same purified siRNA provides three peaks when analyzed by capillary gel electrophoresis (CGE), one peak presumably corresponding to the duplex siRNA,

and two peaks presumably corresponding to the separate siRNA sequence strands. Ion exchange HPLC analysis of the same siRNA contract only shows a single peak. Testing of the purified siRNA construct using a luciferase reporter assay described below demonstrated the same RNAi activity compared to siRNA constructs generated from separately synthesized oligonucleotide sequence strands.

Example 2: Serum stability of chemically modified siRNA constructs

Chemical modifications were introduced into siRNA constructs to determine the stability of these constructs compared to native siRNA oligonucleotides (containing two thymidine nucleotide overhangs) in human serum. An investigation of the serum stability of RNA duplexes revealed that siRNA constructs consisting of all RNA nucleotides containing two thymidine nucleotide overhangs have a half-life in serum of 15 seconds, whereas chemically modified siRNA constructs remained stable in serum for 1 to 3 days depending on the extent of modification. RNAi stability tests were performed by internally labeling one strand (strand 1) of siRNA and duplexing with 1.5 X the concentration of the complimentary siRNA strand (strand 2) (to insure all labeled material was in duplex form). Duplexed siRNA constructs were then tested for stability by incubating at a final concentration of 2μM siRNA (strand 2 concentration) in 90% mouse or human serum for time-points of 30sec, 1min, 5min, 30min, 90min, 4hrs 10min, 16hrs 24min, and 49hrs. Time points were run on a 15% denaturing polyacrylamide gels and analyzed on a phosphoimager.

Internal labeling was performed via kinase reactions with polynucleotide kinase (PNK) and ^{32}P-γ-ATP, with addition of radiolabeled phosphate at nucleotide 13 of strand 2, counting in from the 3' side. Ligation of the remaining 8-mer fragments with T4 RNA ligase resulted in the full length, 21-mer, strand 2. Duplexing of RNAi was done by adding appropriate concentrations of the siRNA oligonucleotides and heating to 95° C for 5min followed by slow cooling to room temperature. Reactions were performed by adding 100% serum to the siRNA duplexes and incubating at 37° C, then removing aliquots at desired time-points. Results of this study are summarized in Figure 3.

Example 3: RNAi activity of chemically modified siRNA constructs

Short interfering RNA (siRNA) is emerging as a powerful tool for gene regulation. All-ribose siRNA duplexes activate the RNAi pathway but have limited utility as therapeutic compounds due to their nuclease sensitivity and short half-life in serum, as shown in Example 2 above. To develop nuclease-resistant siRNA constructs for *in vivo* applications, siRNAs that

target luciferase mRNA and contain stabilizing chemical modifications were tested for activity in HeLa cells. The sequences for the siRNA oligonucleotide sequences used in this study are shown in Table I. Modifications included phosphorothioate linkages (P=S), 2'–*O*–methyl nucleotides, or 2'–fluoro (*F*) nucleotides in one or both siRNA strands. Generally, the sense strand was more sensitive to P=S modifications, whereas the antisense strand was more sensitive to 2'–*O*–methyl modifications. For example, siRNAs containing 10 or 14 P=S at the 5'–end were active when present in the antisense strand and not the sense strand. siRNAs containing 5 or 10 2'–*O*–methyl nucleotides at the 5'–end or 2'–*O*–methyl pyrimidines throughout were active when present in the sense strand and not the antisense strand. 2'–F pyrimidines were however tolerated in either strand alone but not both when both strands are completely substituted. Active siRNA containing stabilizing modifications such as described herein should prove useful for *in vivo* applications.

A luciferase reporter system was utilized to test RNAi activity of chemically modified siRNA constructs compared to siRNA constructs consisting of all RNA nucleotides containing two thymidine nucleotide overhangs. Sense and antisense siRNA strands (20 uM each) were annealed by incubation in buffer (100 mM potassium acetate, 30 mM HEPES-KOH, pH 7.4, 2 mM magnesium acetate) for 1 min. at 90°C followed by 1 hour at 37°C. Plasmids encoding firefly luciferase (pGL2) and renilla luciferase (pRLSV40) were purchased from Promega Biotech.

HeLa S3 cells were grown at 37C in DMEM with 5% FBS and seeded at 15,300 cells in 100 ul media per well of a 96-well plate 24 hours prior to transfection. For transfection, 4 ul Lipofectamine 2000 (Life Technologies) was added to 96 ul OPTI-MEM, vortexed and incubated at room temperature for 5 minutes. The 100 ul diluted lipid was then added to a microtiter tube containing 5 ul pGL2 (200ng/ul), 5 ul pRLSV40 (8 ng/ul) 6 ul siRNA (25 nM or 10 nM final), and 84 ul OPTI-MEM, vortexed briefly and incubated at room temperature for 20 minutes. The transfection mix was then mixed briefly and 50 ul was added to each of three wells that contained HeLa S3 cells in 100 ul media. Cells were incubated for 20 hours after transfection and analyzed for luciferase expression using the Dual luciferase assay according to the manufacturer's instructions (Promega Biotech). The results of this study are summarized in Figures 4-9. The sequences of the siRNA strands used in this study are shown in Table I and are referred to by RPI# in the figures. Normalized luciferase activity is reported as the ratio of firefly luciferase activity to renilla luciferase activity in the same sample. Error bars represent standard deviation of triplicate transfections. As shown in Figures 4-9, the RNAi activity of chemically modified constructs is often comparable to that of control siRNA constructs

consisting of all RNA nucleotides containing two thymidine nucleotide overhangs. In some instances, the RNAi activity of the chemically modified constructs is even better than the siRNA construct consisting of all RNA nucleotides containing two thymidine nucleotide overhangs. For example, Figure 4 shows results obtained from a screen using phosphorothioate modified siRNA constructs; the 27654/27659 construct contains phosphorothioate substitutions for every pyrimidine nucleotide in both sequences, the 27657/27662 construct contains 5 terminal 3'-phosphorothioate substitutions in each strand, the 27649/27658 construct contains all phosphorothioate substitutions only in the antisense strand, whereas the 27649/27660 and 27649/27661 constructs have unmodified sense strands and varying degrees of phosphorothioate substitutions in the antisense strand. All of these constructs show significant RNAi activity when compared to a scrambled siRNA and only slight inhibition of activity when compared to the control siRNA construct.

Figure 5 shows results obtained from a screen using phosphorothioate (28253/28255 and 28254/28256) and universal base substitutions (28257/28259 and 28258/28260) compared to the same controls described above, these modifications show equivalent or better RNAi activity when compared to the control siRNA construct. Figure 6 shows results obtained from a screen using 2'-O-methyl modified siRNA constructs in which the sense strand contains either 10 (28244/27650) or 5 (28245/27650) 2'-O-methyl substitutions, both with comparable activity to the control siRNA construct. Figure 7 shows results obtained from a screen using 2'-O-methyl or 2'-deoxy-2'-fluoro modified siRNA constructs compared to a control construct consisting of all RNA nucleotides containing two thymidine nucleotide overhangs. Figure 8 compares a siRNA construct containing six phosphorothioate substitutions in each strand (28460/28461), where 5 phosphorothioates are present at the 3' end and a single phosphorothioate is present at the 5' end of each strand. This motif shows very similar activity to the control siRNA construct consisting of all RNA nucleotides containing two thymidine nucleotide overhangs. Figure 9 compares a siRNA construct synthesized by the method of the invention described in Example 1, wherein an inverted deoxyabasic succinate linker was used to generate a siRNA having a 3'-inverted deoxyabasic cap on the antisense strand of the siRNA. This construct shows improved activity compared to the control siRNA construct consisting of all RNA nucleotides containing two thymidine nucleotide overhangs.

Example 4: RNAi activity titration

A titration assay was performed to determine the lower range of siRNA concentration required for RNAi activity both in a control siRNA construct consisting of all RNA nucleotides

containing two thymidine nucleotide overhangs and a chemically modified siRNA construct comprising 5 phosphorothioate internucleotide linkages in both the sense and antisense strands. The assay was performed as described in Example 3 above, however, the siRNA constructs were diluted to final concentrations between 2.5 nM and 0.025 nM. Results are shown in Figure 10. As shown in Figure 10, the chemically modified siRNA construct shows a very similar concentration dependent RNAi activity profile to the control siRNA construct when compared to an inverted siRNA sequence control.

Example 5: Identification of potential siRNA target sites in any RNA sequence

The sequence of an RNA target of interest, such as a viral or human mRNA transcript, is screened for target sites, for example by using a computer folding algorithm. In a non-limiting example, the sequence of a gene or RNA gene transcript derived from a database, such as Genbank, is used to generate siRNA targets having complimentarity to the target. Such sequences can be obtained from a database, or can be determined experimentally as known in the art. Target sites that are known, for example, those target sites determined to be effective target sites based on studies with other nucleic acid molecules, for example ribozymes or antisense, or those targets known to be associated with a disease or condition such as those sites containing mutations or deletions, can be used to design siRNA molecules targeting those sites as well. Various parameters can be used to determine which sites are the most suitable target sites within the target RNA sequence. These parameters include but are not limited to secondary or tertiary RNA structure, the nucleotide base composition of the target sequence, the degree of homology between various regions of the target sequence, or the relative position of the target sequence within the RNA transcript. Based on these determinations, any number of target sites within the RNA transcript can be chosen to screen siRNA molecules for efficacy, for example by using in vitro RNA cleavage assays, cell culture, or animal models. In a non-limiting example, anywhere from 1 to 1000 target sites are chosen within the transcript based on the size of the siRNA contruct to be used. High throughput screening assays can be developed for screening siRNA molecules using methods known in the art, such as with multi-well or multi-plate assays to determine efficient reduction in target gene expression.

Example 6: Selection of siRNA molecule target sites in a RNA

siRNA target sites were chosen by analyzing sequences of the RNA target and optionally prioritizing the target sites on the basis of folding (structure of any given sequence analyzed to determine siRNA accessibility to the target). siRNA molecules are designed that could bind each target and are optionally individually analyzed by computer folding to assess whether the siRNA molecule can interact with the target sequence. Varying the length of the siRNA molecules can be chosen to optimize activity. Generally, a sufficient number of complimentary nucleotide bases are chosen to bind to, or otherwise interact with, the target RNA, but the degree of complementarity can be modulated to accommodate siRNA duplexes or varying length or base composition. By using such methodologies, siRNA molecules can be designed to target sites within any known RNA sequence, for example those RNA sequences corresponding to the any gene transcript.

Example 7: Chemical Synthesis and Purification of siRNA

siRNA molecules can be designed to interact with various sites in the RNA message, for example target sequences within the RNA sequences described herein. The sequence of one strand of the siRNA molecule(s) are complementary to the target site sequences described above. The siRNA molecules can be chemically synthesized using methods described herein. Inactive siRNA molecules that are used as control sequences can be synthesized by scrambling the sequence of the siRNA molecules such that it is not complimentary to the target sequence.

Example 8: siRNA Inhibition of a chimeric HCV/Poliovirus in HeLa Cells

Inhibition of a chimeric HCV/Poliovirus was investigated using 21 nucleotide siRNA duplexes in HeLa cells. Seven siRNA were designed that target three regions in the highly conserved 5' untranslated region (UTR) of HCV RNA. The siRNAs were screened in two cell culture systems dependent upon the 5'-UTR of HCV; one requires translation of an HCV/luciferase gene, while the other involves replication of a chimeric HCV/poliovirus (PV) (see Blatt *et al.*, USSN 09/740,332, filed December 18, 2000, incorporated by reference herein). Two siRNAs targeting the same region (shifted by one nucleotide) are active in both systems (see Figure 11). For example, a >85% reduction in HCVPV replication was observed in siRNA-

treated cells compared to a inverse siRNA control (Figure 11) with an IC50 = ~2.5 nM (Figure 12) using SED ID NOS: 198 and 205 as the siRNA sense and antisense strands respectively. To develop nuclease-resistant siRNA for in vivo applications, siRNAs can be modified to contain stabilizing chemical modifications. Such modifications include phosphorothioate linkages (P=S), 2'-*O*-methyl nucleotides, 2'-fluoro (F) nucleotides, 5' and/or 3' end modifications and a variety of other nucleotide and non-nucleotide modifications, in one or both siRNA strands. Using this systematic approach, active siRNA molecules have been identified that are substantially more resistant to nucleases (see Example 2). Two of these constructs were tested in the HCV/poliovirus chimera system, both demonstrating significant reduction in viral replication (Figures 13 and 14). As such, chemically modified, nuclease resistant siRNA molecules represent an important class of therapeutic agents for treating chronic HCV infection.

Animal Models

Various animal models can be used to screen siRNA constructs *in vivo* as are known in the art, for example those animal models that are used to evaluate other nucleic acid technologies such as enzymatic nucleic acid molecules (ribozymes) and/or antisense. Such animal models are used to test the efficacy of siRNA molecules described herein. In a non-limiting example, siRNA molecules that are designed as anti-angiogenic agents can be screened animal models.. There are several animal models in which the anti-angiogenesis effect of nucleic acids of the present invention, such as siRNA, directed against genes associated with angiogenesis and/or metastais, such as VEGF-R genes. Typically a corneal model has been used to study angiogenesis in rat and rabbit since recruitment of vessels can easily be followed in this normally avascular tissue (Pandey *et al.*, 1995 *Science* 268: 567-569). In these models, a small Teflon or Hydron disk pretreated with an angiogenesis factor (e.g. bFGF or VEGF) is inserted into a pocket surgically created in the cornea. Angiogenesis is monitored 3 to 5 days later. siRNA molecules directed against VEGF-R mRNAs would be delivered in the disk as well, or dropwise to the eye over the time course of the experiment. In another eye model, hypoxia has been shown to cause both increased expression of VEGF and neovascularization in the retina (Pierce *et al.*, 1995 *Proc. Natl. Acad. Sci.* USA. 92: 905-909; Shweiki *et al.*, 1992 *J. Clin. Invest.* 91: 2235-2243).

Several animal models exist for screening of anti-angiogenic agents. These include corneal vessel formation following corneal injury (Burger *et al.*, 1985 *Cornea* 4: 35-41; Lepri, *et al.*, 1994 *J. Ocular Pharmacol.* 10: 273-280; Ormerod *et al.*, 1990 *Am. J. Pathol.* 137:

1243-1252) or intracorneal growth factor implant (Grant *et al.*, 1993 *Diabetologia* 36: 282-291; Pandey *et al.* 1995 *supra;* Zieche *et al.*, 1992 *Lab. Invest.* 67: 711-715), vessel growth into Matrigel matrix containing growth factors (Passaniti *et al.*, 1992 *supra*), female reproductive organ neovascularization following hormonal manipulation (Shweiki *et al.*, 1993 *Clin. Invest.* 91: 2235-2243), several models involving inhibition of tumor growth in highly vascularized solid tumors (O'Reilly *et al.*, 1994 *Cell* 79: 315-328; Senger *et al.*, 1993 *Cancer and Metas. Rev.* 12: 303-324; Takahasi *et al.*, 1994 *Cancer Res.* 54: 4233-4237; Kim *et al.*, 1993 *supra*), and transient hypoxia-induced neovascularization in the mouse retina (Pierce *et al.*, 1995 *Proc. Natl. Acad. Sci.* USA. 92: 905-909).

The cornea model, described in Pandey et al. *supra*, is the most common and well characterized anti-angiogenic agent efficacy screening model. This model involves an avascular tissue into which vessels are recruited by a stimulating agent (growth factor, thermal or alkalai burn, endotoxin). The corneal model would utilize the intrastromal corneal implantation of a Teflon pellet soaked in a VEGF-Hydron solution to recruit blood vessels toward the pellet which can be quantitated using standard microscopic and image analysis techniques. To evaluate their anti-angiogenic efficacy, ribozymes are applied topically to the eye or bound within Hydron on the Teflon pellet itself. This avascular cornea as well as the Matrigel model provide for low background assays. While the corneal model has been performed extensively in the rabbit, studies in the rat have also been conducted.

The mouse model (Passaniti et al., *supra*) is a non-tissue model which utilizes Matrigel, an extract of basement membrane (Kleinman et al., 1986) or Millipore® filter disk, which can be impregnated with growth factors and anti-angiogenic agents in a liquid form prior to injection. Upon subcutaneous administration at body temperature, the Matrigel or Millipore® filter disk forms a solid implant. VEGF embedded in the Matrigel or Millipore® filter disk would be used to recruit vessels within the matrix of the Matrigel or Millipore® filter disk which can be processed histologically for endothelial cell specific vWF (factor VIII antigen) immunohistochemistry, Trichrome-Masson stain, or hemoglobin content. Like the cornea, the Matrigel or Millipore® filter disk are avascular; however, it is not tissue. In the Matrigel or Millipore® filter disk model, siRNA molecules are administered within the matrix of the Matrigel or Millipore® filter disk to test their anti-angiogenic efficacy. Thus, delivery issues in this model, as with delivery of siRNA molecules by Hydron- coated Teflon pellets in the rat cornea model, may be less problematic due to the homogeneous presence of the siRNA within the respective matrix.

The Lewis lung carcinoma and B-16 murine melanoma models are well accepted models of primary and metastatic cancer and are used for initial screening of anti-cancer agents. These murine models are not dependent upon the use of immunodeficient mice, are relatively inexpensive, and minimize housing concerns. Both the Lewis lung and B-16 melanoma models involve subcutaneous implantation of approximately 10^6 tumor cells from metastatically aggressive tumor cell lines (Lewis lung lines 3LL or D122, LLc-LN7; B-16-BL6 melanoma) in C57BL/6J mice. Alternatively, the Lewis lung model can be produced by the surgical implantation of tumor spheres (approximately 0.8 mm in diameter). Metastasis also may be modeled by injecting the tumor cells directly *i.v.*. In the Lewis lung model, microscopic metastases can be observed approximately 14 days following implantation with quantifiable macroscopic metastatic tumors developing within 21-25 days. The B-16 melanoma exhibits a similar time course with tumor neovascularization beginning 4 days following implantation. Since both primary and metastatic tumors exist in these models after 21-25 days in the same animal, multiple measurements can be taken as indices of efficacy. Primary tumor volume and growth latency as well as the number of micro- and macroscopic metastatic lung foci or number of animals exhibiting metastases can be quantitated. The percent increase in lifespan can also be measured. Thus, these models would provide suitable primary efficacy assays for screening systemically administered siRNA molecules and siRNA formulations.

In the Lewis lung and B-16 melanoma models, systemic pharmacotherapy with a wide variety of agents usually begins 1-7 days following tumor implantation/inoculation with either continuous or multiple administration regimens. Concurrent pharmacokinetic studies can be performed to determine whether sufficient tissue levels of siRNA can be achieved for pharmacodynamic effect to be expected. Furthermore, primary tumors and secondary lung metastases can be removed and subjected to a variety of *in vitro* studies (i.e. target RNA reduction).

In utilizing these models to assess siRNA activity, VEGFR1 and/or VEGFR2 protein levels can be measured clinically or experimentally by FACS analysis. VEGFR1 and/or VEGFR2 encoded mRNA levels will be assessed by Northern analysis, RNase-protection, primer extension analysis and/or quantitative RT-PCR. siRNA molecules that block VEGFR1 and/or VEGFR2 protein encoding mRNAs and therefore result in decreased levels of VEGFR1 and/or VEGFR2 activity by more than 20% *in vitro* can be thus identified.

Indications

The siRNA molecules of the invention can be used to treat a variety of diseases and conditions through modulation of gene expression. Using the methods described herein, chemically modified siRNA molecules can be designed to modulate the expression any number of target genes, including but not limited to genes associated with cancer, metabolic diseases, infectious diseases such as viral, bacterial or fungal infections, neurologic diseases, musculoskeletal diseases, diseases of the immune system, diseases associated with signaling pathways and cellular messengers, and diseases associated with transport systems including molecular pumps and channels.

Non-limiting examples of various viral genes that can be targeted using siRNA molecules of the invention include Hepatitis C Virus (HCV, for example Genbank Accession Nos: D11168 , D50483.1, L38318 and S82227), Hepatitis B Virus (HBV, for example GenBank Accession No. AF100308.1), Human Immunodeficiency Virus type 1 (HIV-1, for example GenBank Accession No. U51188), Human Immunodeficiency Virus type 2 (HIV-2, for example GenBank Accession No. X60667), West Nile Virus (WNV for example GenBank accession No. NC_001563), cytomegalovirus (CMV for example GenBank Accession No. NC_001347), respiratory syncytial virus (RSV for example GenBank Accession No. NC_001781), influenza virus (for example example GenBank Accession No. AF037412, rhinovirus (for example, GenBank accession numbers: D00239, X02316, X01087, L24917, M16248, K02121, X01087), papillomavirus (for example GenBank Accession No. NC_001353), Herpes Simplex Virus (HSV for example GenBank Accession No. NC_001345), and other viruses such as HTLV (for example GenBank Accession No. AJ430458). Due to the high sequence variability of many viral genomes, selection of siRNA molecules for broad therapeutic applications would likely involve the conserved regions of the viral genome. Nonlimiting examples of conserved regions of the viral genomes include but are not limited to 5'-Non Coding Regions (NCR), 3'- Non Coding Regions (NCR) and/or internal ribosome entry sites (IRES). siRNA molecules designed against conserved regions of various viral genomes will enable efficient inhibition of viral replication in diverse patient populations and may ensure the effectiveness of the siRNA molecules against viral quasi species which evolve due to mutations in the non-conserved regions of the viral genome.

Non-limiting examples of human genes that can be targeted using siRNA molecules of the invention using methods described herein include any human RNA sequence, for example those commonly referred to by Genbank Accession Number. These RNA sequences can be used to design siRNA molecules that inhibit gene expression and therefore abrogate diseases, conditions, or infections associated with expression of those genes. Such non-limiting examples of human

genes that can be targeted using siRNA molecules of the invention include VEGFr (VEGFr-1 for example GenBank Accession No. XM_067723, VEGFr-2 for example GenBank Accession No. AF063658), HER1, HER2, HER3, and HER4 (for example Genbank Accession Nos: NM_005228, NM_004448, NM_001982, and NM_005235 respectively), telomerase (TERT, for example GenBank Accession No. NM_003219), telomerase RNA (for example GenBank Accession No. U86046), NFkappaB, Rel-A (for example GenBank Accession No. NM_005228), NOGO (for example GenBank Accession No. AB020693), NOGOr (for example GenBank Accession No. XM_015620), RAS (for example GenBank Accession No. NM_004283), RAF (for example GenBank Accession No. XM_033884), CD20 (for example GenBank Accession No. X07203), METAP2 (for example GenBank Accession No. NM_003219), CLCA1 (for example GenBank Accession No. NM_001285) , phospholamban (for example GenBank Accession No. NM_002667), PTP1B (for example GenBank Accession No. M31724), and others.

The siRNA molecule of the invention can also be used in a variety of agricultural applications involving modulation of endogenous or exogenous gene expression in plants using siRNA, including use as insecticidal, antiviral and anti-fungal agents or modulate plant traits such as oil and starch profiles and stress resistance.

Diagnostic uses

The siRNA molecules of the invention can be used in a variety of diagnostic applications, such as in identifying molecular targets such as RNA in a variety of applications, for example, in clinical, industrial, environmental, agricultural and/or research settings. Such diagnostic use of siRNA molecules involves utilizing reconstituted RNAi systems, for example using cellular lysates or partially purified cellular lysates. siRNA molecules of this invention may be used as diagnostic tools to examine genetic drift and mutations within diseased cells or to detect the presence of endogenous or exogenous, for example viral, RNA in a cell. The close relationship between siRNA activity and the structure of the target RNA allows the detection of mutations in any region of the molecule, which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple siRNA molecules described in this invention, one may map nucleotide changes, which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with siRNA molecules can be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease or infection. In this manner, other genetic targets may be defined as important mediators

of the disease. These experiments will lead to better treatment of the disease progression by affording the possibility of combination therapies (e.g., multiple siRNA molecules targeted to different genes, siRNA molecules coupled with known small molecule inhibitors, or intermittent treatment with combinations siRNA molecules and/or other chemical or biological molecules). Other *in vitro* uses of siRNA molecules of this invention are well known in the art, and include detection of the presence of mRNAs associated with a disease, infection, or related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a siRNA using standard methodologies, for example fluorescence resonance emission transfer (FRET).

In a specific example, siRNA molecules that can cleave only wild-type or mutant forms of the target RNA are used for the assay. The first siRNA molecules is used to identify wild-type RNA present in the sample and the second siRNA molecules will be used to identify mutant RNA in the sample. As reaction controls, synthetic substrates of both wild-type and mutant RNA will be cleaved by both siRNA molecules to demonstrate the relative siRNA efficiencies in the reactions and the absence of cleavage of the "non-targeted" RNA species. The cleavage products from the synthetic substrates will also serve to generate size markers for the analysis of wild-type and mutant RNAs in the sample population. Thus each analysis will require two siRNA molecules, two substrates and one unknown sample which will be combined into six reactions. The presence of cleavage products will be determined using an RNase protection assay so that full-length and cleavage fragments of each RNA can be analyzed in one lane of a polyacrylamide gel. It is not absolutely required to quantify the results to gain insight into the expression of mutant RNAs and putative risk of the desired phenotypic changes in target cells. The expression of mRNA whose protein product is implicated in the development of the phenotype (*i.e.*, disease related or infection related) is adequate to establish risk. If probes of comparable specific activity are used for both transcripts, then a qualitative comparison of RNA levels will be adequate and will decrease the cost of the initial diagnosis. Higher mutant form to wild-type ratios will be correlated with higher risk whether RNA levels are compared qualitatively or quantitatively.

All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains. All references cited in this

disclosure are incorporated by reference to the same extent as if each reference had been incorporated by reference in its entirety individually.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The methods and compositions described herein as presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art, which are encompassed within the spirit of the invention, are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. Thus, such additional embodiments are within the scope of the present invention and the following claims.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations that are not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments, optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the description and the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

Table I

RPI#	Aliases	Sequence	SEQ ID#
25227	RPI 21550 EGFR 3830L23 AS as siRNA Str 1 (sense)	B UAACCUCGUACUGGUGCCUCC B	1
25228	RPI 21550 EGFR 3830L23 AS as siRNA Str 2 (antisense)	B GGAGGCACCAGUACGAGGUUA B	2
25229	RPI 21549 EGFR as siRNA Str 2 (antisense)	B AAACUCCAAGAUCCCCAAUCA B	3
25230	RPI 21549 EGFR 3 as siRNA Str 1 (sense)	B UGAUUGGGGAUCUUGGAGUUU B	4
25231	RPI 21547EGFR as siRNA Str 2 (antisense)	B GUUGGAGUCUGUAGGACUUGG B	5
25232	RPI 21547EGFR as siRNA Str 1 (sense)	B CCAAGUCCUACAGACUCCAAC B	6
25233	RPI 21545 EGFR as siRNA Str 2 (antisense)	B GCAAAAACCCUGUGAUUUCCU B	7
25234	RPI 21545 EGFR as siRNA Str 1 (sense)	B AGGAAAUCACAGGGUUUUUGC B	8
25235	RPI 21543 EGFR as siRNA Str 2 (antisense)	B UUGGUCAGUUUCUGGCAGUUC B	9
25236	RPI 21543 EGFR as siRNA Str 1 (sense)	B GAACUGCCAGAAACUGACCAA B	10
25237	HCV IRES Loop IIIb (Heptazyme site) as siRNA str1 (sense)	B GGUCCUUUCUUGGAUCAACCC B	11
25238	HCV IRES Loop IIIb (Heptazyme site) as siRNA str2 (antisense)	B GGGUUGAUCCAAGAAAGGACC B	12
25239	HBV (HepBzyme site) as siRNA str1(sense)	B UGGACUUCUCUCAAUUUUCUA B	13
25240	HBV (HepBzyme site) as siRNA str2 (antisense)	B UAGAAAAUUGAGAGAAGUCCA B	14
25241	HBV16371 site as siRNA str1(sense)	B UUUUUCACCUCUGCCUAAUCA B	15
25242	HBV16371 site as siRNA str2 (antisense)	B UGAUUAGGCAGAGGUGAAAAA B	16
25243	HBV16372-16373 site as siRNA str1(sense)	B CAAGCCUCCAAGCUGUGCCUU B	17
25244	HBV16372-16373 site as siRNA str 2 (antisense)	B AAGGCACAGCUUGGAGGCUUG B	18
25245	RPI 17783 Her2Neu AS as siRNA Str 2 (antisense)	B UCCAUGGUGCUCACUGCGGCU B	19
25246	RPI 17783 Her2Neu AS as siRNA Str 1 (sense)	B AGCCGCAGUGAGCACCAUGGA B	20
25247	RPI 17783 Her2Neu AS as siRNA Str 1 (sense) Inverted control	B AGGUACCACGAGUGACGCCGA B	21
25248	RPI 17783 Her2Neu AS as siRNA Str 1 (sense) Inverted control compliment	B UCGGCGUCACUCGUGGUACCU B	22
25249	RPI 21550 EGFR 3830L23 AS as siRNA Str 1 (sence) Inverted Control	B CCUCCGUGGUCAUGCUCCAAU B	23
25250	RPI 21650 EGFR 3830L23 AS as siRNA Str 1 (sence) Inverted Control Compliment	B AUUGGAGCAUGACCACGGAGG B	24
25251	HCV IRES Loop IIIb (Heptazyme site) as siRNA str1 (sense) Inverted Control	B CCCAACUAGGUUCUUUCCUGG B	25
25252	HCV IRES Loop IIIb (Heptazyme site) as siRNA str1 (sense) Inverted	B CCAGGAAAGAACCUAGUUGGG B	26

	Control Compliment		
25804	RPI 21550 EGFR 3830L23 AS as siRNA Str 1 (sense) +2U overhang	UAACCUCGUACUGGUGCCUCCUU	27
25805	RPI 21550 EGFR 3830L23 AS as siRNA Str 2 (antisense) +2U overhang	GGAGGCACCAGUACGAGGUUAUU	28
25808	RPI 21549 EGFR as siRNA Str 2 (antisense)+ 2U overhang	AAACUCCAAGAUCCCCAAUCAUU	29
25824	RPI 21550 EGFR 3830L23 AS as siRNA Str 1 (sense) +2U overhang	BUAACCUCGUACUGGUGCCUCCUUB	30
25825	RPI 21550 EGFR 3830L23 AS as siRNA Str 2 (antisense) +2U overhang	BGGAGGCACCAGUACGAGGUUAUUB	31
25826	RPI 21549 EGFR as siRNA Str 2 (antisense)+ 2U overhang	BAAACUCCAAGAUCCCCAAUCAUUB	32
25807	RPI 21549 EGFR 3 as siRNA Str 1 (sense)+2U overhang	UGAUUGGGGAUCUUGGAGUUUUU	33
25808	RPI 21547EGFR as siRNA Str 2 (antisense) +2U overhang	GUUGGAGUCUGUAGGACUUGGUU	34
25809	RPI 21547EGFR as siRNA Str 1 (sense) + 2U overhang	CCAAGUCCUACAGACUCCAACUU	35
25827	RPI 21549 EGFR 3 as siRNA Str 1 (sense)+2U overhang	BUGAUUGGGGAUCUUGGAGUUUUUB	36
25828	RPI 21547EGFR as siRNA Str 2 (antisense) +2U overhang	BGUUGGAGUCUGUAGGACUUGGUUB	37
25829	RPI 21547EGFR as siRNA Str 1 (sense) + 2U overhang	BCCAAGUCCUACAGACUCCAACUUB	38
25810	RPI 21545 EGFR as siRNA Str 2 (antisense)+2U overhang	GCAAAAACCCUGUGAUUUCCUUU	39
25811	RPI 21545 EGFR as siRNA Str 1 (sense)+2U overhang	AGGAAAUCACAGGGUUUUUGCUU	40
25812	RPI 21543 EGFR as siRNA Str 2 (antisense)+2U overhang	UUGGUCAGUUUCUGGCAGUUCUU	41
25830	RPI 21545 EGFR as siRNA Str 2 (antisense)+2U overhang	BGCAAAAACCCUGUGAUUUCCUUUB	42
25831	RPI 21545 EGFR as siRNA Str 1 (sense)+2U overhang	BAGGAAAUCACAGGGUUUUUGCUUB	43
25832	RPI 21543 EGFR as siRNA Str 2 (antisense)+2U overhang	BUUGGUCAGUUUCUGGCAGUUCUUB	44
25813	RPI 21543 EGFR as siRNA Str 1 (sense)+2U overhang	GAACUGCCAGAAACUGACCAAUU	45
25814	HCV IRES Loop IIIb (Heptazyme site) as siRNA str1 (sense)+2U overhang	GGUCCUUUCUUGGAUCAACCCUU	46
25816	HCV IRES Loop IIIb (Heptazyme site) as siRNA str2 (antisense) +2U overhang	GGGUUGAUCCAAGAAAGGACCUU	47
25833	RPI 21543 EGFR as siRNA Str 1 (sense)+2U overhang	BGAACUGCCAGAAACUGACCAAUUB	48
25834	HCV IRES Loop IIIb (Heptazyme site) as siRNA str1 (sense)+2U overhang	BGGUCCUUUCUUGGAUCAACCCUUB	49
25835	HCV IRES Loop IIIb (Heptazyme site) as siRNA str2 (antisense) +2U overhang	BGGGUUGAUCCAAGAAAGGACCUUB	50
25818	HBV (HepBzyme site) as siRNA str1(sense)+2U overhang	UGGACUUCUCUCAAUUUUCUAUU	51
25817	HBV (HepBzyme site) as siRNA str2 (antisense)+2U overhang	UAGAAAAUUGAGAGAAGUCCAUU	52
25818	HBV16371 site as siRNA	UUUUUCACCUCUGCCUAAUCAUU	53

	str1(sense)+2U overhang		
25836	HBV (HepBzyme site) as siRNA str1(sense)+2U overhang	BUGGACUUCUCUCAAUUUUCUAUUB	54
25837	HBV (HepBzyme site) as siRNA str2 (antisense)+2U overhang	BUAGAAAAUUGAGAGAAGUCCAUUB	55
25838	HBV18371 site as siRNA str1(sense)+2U overhang	BUUUUUCACCUCUGCCUAAUCAUUB	56
25819	HBV18371 site as siRNA str2 (antisense)+2U overhang	UGAUUAGGCAGAGGUGAAAAAUU	57
25820	HBV18372-18373 site as siRNA str1(sense)+2U overhang	CAAGCCUCCAAGCUGUGCCUUUU	58
25821	HBV18372-18373 site as siRNA str 2 (antisense)+2U overhang	AAGGCACAGCUUGGAGGCUUGUU	59
25839	HBV18371 site as siRNA str2 (antisense)+2U overhang	BUGAUUAGGCAGAGGUGAAAAAUUB	60
25840	HBV18372-18373 site as siRNA str1(sense)+2U overhang	BCAAGCCUCCAAGCUGUGCCUUUUB	61
25841	HBV18372-18373 site as siRNA str 2 (antisense)+2U overhang	BAAGGCACAGCUUGGAGGCUUGUUB	62
25822	RPI 17763 Her2Neu AS as siRNA Str 2 (antisense)+2U overhang	UCCAUGGUGCUCACUGCGGCUUU	63
25823	RPI 17763 Her2Neu AS as siRNA Str 1 (sense)+2U overhang	AGCCGCAGUGAGCACCAUGGAUU	64
25842	RPI 17763 Her2Neu AS as siRNA Str 2 (antisense)+2U overhang	BUCCAUGGUGCUCACUGCGGCUUUB	65
25843	RPI 17763 Her2Neu AS as siRNA Str 1 (sense)+2U overhang	BAGCCGCAGUGAGCACCAUGGAUUB	66
27649	RPI GL2 Str1	CGUACGCGGAAUACUUCGA TT	67
27650	RPI GL2 Str2	UCGAAGUAUUCCGCGUACG TT	68
27651	RPI Inverted GL2 Str1	AGCUUCAUAAGGCGCAUGC TT	69
27652	RPI Inverted GL2 Str2	GCAUGCGCCUUAUGAAGCU TT	70
27653	RPI GL2 Str1 all ribo P=S	$C_sG_sU_sA_sC_sG_sC_sG_sG_sA_sA_sU_sA_sC_sU_sU_sC_sG_sA$ TT	71
27654	RPI GL2 Str1 all ribo pyrimidines P=S	$C_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGA$ TT	72
27655	RPI GL2 Str1 14 5' P=S	$C_sG_sU_sA_sC_sG_sC_sG_sG_sA_sA_sU_sA_sC_sUUCGA$ TT	73
27656	RPI GL2 Str1 10 5' P=S	$C_sG_sU_sA_sC_sG_sC_sG_sG_sA_sAUACUUCGA$ TT	74
27657	RPI GL2 Str1 5 5' P=S	$C_sG_sU_sA_sC_sGCGGAAUACUUCGA$ TT	75
27658	RPI GL2 Str2 all ribo P=S	$U_sC_sG_sA_sA_sG_sU_sA_sU_sU_sC_sC_sG_sC_sG_sU_sA_sC_sG$ TT	76
27659	RPI GL2 Str2 all ribo pyrimidines P=S	$U_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sG$ TT	77
27660	RPI GL2 Str2 5' 14 P=S	$U_sC_sG_sA_sA_sGU_sA_sU_sU_sC_sC_sG_sC_sGUACG$ TT	78
27661	RPI GL2 Str2 5' 10 P=S	$U_sC_sG_sA_sA_sG_sU_sA_sU_sU_sCCGCGUACG$ TT	79
27662	RPI GL2 Str2 5' 5 P=S	$U_sC_sG_sA_sA_sGUAUUCCGCGUACG$ TT	80
28010	RPI GL2 Str1 5'ligation fragment	CGUACG	81
28011	RPI GL2 Str1 3' ligation fragment	CGGAAUACUUCGATT	82
28012	RPI GL2 Str2 5' ligation fragment	UCGAAGUA	83
28013	RPI GL2 Str2 3'ligation fragment	UUCCGCGUACGTT	84
28254	RPI GL2 Str1 all pyrimidines + TT = PS	$C_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGAT_sT$	85
28255	RPI GL2 Str2, + TT = PS	$UCGAAGUAUUCCGCGUACGT_sT$	86
28256	RPI GL2 Str2, all pyrimidines+ TT = PS	$U_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sGT_sT$	87
28292	Her2.1.sense Str1	UGGGGUCGUCAAAGACGUUTT	88
28283	Her2.1.antisense Str2	AACGUCUUUGACGACCCCATT	89
28284	Her2.1.sense Str1 inverted	UUGCAGAAACUGCUGGGGUTT	90

28265	Her2.1.antisense Str2 inverted	ACCCCAGCAGUUUCUGCAATT	91
28266	Her2.2.sense Str1	GGUGCUUGGAUCUGGCGCUTT	92
28267	Her2.2.antisense Str2	AGCGCCAGAUCCAAGCACCTT	93
28268	Her2.2.sense Str1 inverted	UCGCGGUCUAGGUUCGGUGTT	94
28269	Her2.2.antisense Str2 inverted	CCACGAACCUAGACCGCGATT	95
28270	Her2.3.sense Str1	GAUCUUUGGGAGCCUGGCATT	96
28271	Her2.3.antisense Str2	UGCCAGGCUCCCAAAGAUCTT	97
28272	Her2.3.sense Str1 inverted	ACGGUCCGAGGGUUUCUAGTT	98
28273	Her2.3.antisense Str2 inverted	CUAGAAACCCUCGGACCGUTT	99
28274	RPI inverted GL2 Str1 all ribo pyrimidines P=S	$AGC_sU_sU_sC_sAU_sAAGGC_sGC_sAU_sGC$ TT	100
28275	RPI inverted GL2 Str1 5 5' P=S	$A_sG_sC_sU_sU_sCAUAAGGCGCAUGC$ TT	101
28276	RPI inverted GL2 Str2 all ribo pyrimidines P=S	$GC_sAU_sGC_sGC_sC_sU_sU_sAU_sGAAGC_sU$ TT	102
28277	RPI inverted GL2 Str2 5 5' P=S	$G_sC_sA_sU_sG_sCGCCUUAUGAAGCU$ TT	103
28278	RPI inverted GL2 Str2 all ribo P=S	$G_sC_sA_sU_sG_sC_sG_sC_sC_sU_sU_sA_sU_sG_sA_sA_sG_sC_sU$ TT	104
28279	RPI inverted GL2 Str2 14 5' P=S	$G_sC_sA_sU_sG_sC_sG_sC_sC_sU_sU_sA_sU_sG_sAAGCU$ TT	105
28280	RPI inverted GL2 Str2 10 5' P=S	$G_sC_sA_sU_sG_sC_sG_sC_sC_sU_sUAUGAAGCU$ TT	106
28383	hRelA.1.sense Str1	CAGCACAGACCCAGCUGUGTT	107
28384	hRelA.1.antisense Str2	CACAGCUGGGUCUGUGCUGTT	108
28385	hRelA.1.sense Str1 inverted	GUGUCGACCCAGACACGACTT	109
28386	hRelA.1.antisense Str2 inverted	GUCGUGUCUGGGUCGACACTT	110
28387	hRelA.2.sense Str1	GCAGGCUGGAGGUAAGGCCTT	111
28388	hRelA.2.antisense Str2	GGCCUUACCUCCAGCCUGCTT	112
28389	hRelA.2.sense Str1 inverted	CCGGAAUGGAGGUCGGACGTT	113
28390	hRelA.2.antisense Str2 inverted	CGUCCGACCUCCAUUCCGGTT	114
28391	hmRelA.3.sense Str1	GACUUCUCCUCCAUUGCGGTT	115
28392	hmRelA.3.antisense Str2	CCGCAAUGGAGGAGAAGUCTT	116
28393	hmRelA.3.sense Str1 inverted	GGCGUUACCUCCUCUUCAGTT	117
28394	hmRelA.3.antisense Str2 inverted	CUGAAGAGGAGGUAACGCCTT	118
28395	hmRelA.4.sense Str1	CACUGCCGAGCUCAAGAUCTT	119
28396	hmRelA.4.antisense Str2	GAUCUUGAGCUCGGCAGUGTT	120
28397	hmRelA.4.sense Str1 inverted	CUAGAACUCGAGCCGUCACTT	121
28398	hmRelA.4.antisense Str2 inverted	GUGACGGCUCGAGUUCUAGTT	122
28399	hIKKg.1.sense Str1	GGAGUUCCUCAUGUGCAAGTT	123
28400	hIKKg.1.antisense Str2	CUUGCACAUGAGGAACUCCTT	124
28401	hIKKg.1.sense Str1 inverted	GAACGUGUACUCCUUGAGGTT	125
28402	hIKKg.1.antisense Str2 inverted	CCUCAAGGAGUACACGUUCTT	126
28403	hIKKg.2.sense Str1	UCAAGAGCUCCGAGAUGCCTT	127
28404	hIKKg.2.antisense Str2	GGCAUCUCGGAGCUCUUGATT	128
28405	hIKKg.2.sense Str1 inverted	CCGUAGAGCCUCGAGAACUTT	129
28406	hIKKg.2.antisense Str2 inverted	AGUUCUCGAGGCUCUACGGTT	130
28407	hmIKKG.sense Str1	GCAGAUGGCUGAGGACAAGTT	131
28408	hmIKKG.3.antisense Str2	CUUGUCCUCAGCCAUCUGCTT	132
28409	hmIKKG.3.sense Str1 inverted	GAACAGGAGUCGGUAGACGTT	133
28410	hmIKKG.3.antisense Str2 inverted	CGUCUACCGACUCCUGUUCTT	134
28447	RPI construct as hairpin +GAAA+AU blunt	AACGUACGCGGAAUACUUCGAUUAAAAGUAAUCGG AAGUAUUCCGCGUACGUU	135
28448	RPI construct as hairpin +GAAA+AU 3' overhang	CGUACGCGGAAUACUUCGAUUAAAAGUAAUCGAA GUAUUCCGCGUACGUU	136
28449	RPI construct as hairpin +GAAA blunt	AACGUACGCGGAAUACUUCGAUUAAAGAAUCGAA GUAUUCCGCGUACGUU	137

28450	RPI construct as hairpin +GAAA 3' overhang	CGUACGCGGAAUACUUCGAUUAAAGAAUCGAAG UAUUCCGCGUACGUU	138
28451	RPI construct as hairpin +UUG 3' overhang	CGUACGCGGAAUACUUCGAUUGUUAAUCGAAGU AUUCCGCGUACGUU	139
28452	RPI construct as hairpin +UUG blunt	AACGUACGCGGAAUACUUCGAUUGUUAAUCGAA GUAUUCCGCGUACGUU	140
28453	RPI construct as hairpin +UUG+AU blunt	AACGUACGCGGAAUACUUCGAUUAGUUUAAUCG AAGUAUUCCGCGUACGUU	141
28454	RPI construct as hairpin +UUG 3' overhang	CGUACGCGGAAUACUUCGAUUAGUUUAAUCGAA GUAUUCCGCGUACGUU	142
28415	HCV-Luc:325U21 TT siRNA	CCCCGGGAGGUCUCGUAGATT	143
28416	HCV-Luc:182U21 TT siRNA	CGGAACCGGUGAGUACACCTT	144
28417	HCV-Luc:324U21 TT siRNA	GCCCCGGGAGGUCUCGUAGTT	145
28418	HCV-Luc:183U21 TT siRNA	GGAACCGGUGAGUACACCGTT	146
28419	HCV-Luc:294U21 TT siRNA	GUGGUACUGCCUGAUAGGGTT	147
28420	HCV-Luc:293U21 TT siRNA	UGUGGUACUGCCUGAUAGGTT	148
28421	HCV-Luc:292U21 TT siRNA	UUGUGGUACUGCCUGAUAGTT	149
28422	HCV-Luc:343L21 TT siRNA (325C)	UCUACGAGACCUCCCGGGGTT	150
28423	HCV-Luc:180L21 TT siRNA (182C)	GGUGUACUCACCGGUUCCGTT	151
28424	HCV-Luc:342L21 TT siRNA (324C)	CUACGAGACCUCCCGGGGCTT	152
28425	HCV-Luc:181L21 TT siRNA (183C)	CGGUGUACUCACCGGUUCCTT	153
28426	HCV-Luc:312L21 TT siRNA (294C)	CCCUAUCAGGCAGUACCACTT	154
28427	HCV-Luc:311L21 TT siRNA (293C)	CCUAUCAGGCAGUACCACATT	155
28428	HCV-Luc:310L21 TT siRNA (292C)	CUAUCAGGCAGUACCACAATT	156
28429	HCV-Luc:325U21 TT siRNA Inv	TTAGAUGCUCUGGAGGGCCCC	157
28430	HCV-Luc:182U21 TT siRNA Inv	TTCCACAUCGAGUGGCCAAGGC	158
28431	HCV-Luc:324U21 TT siRNA Inv	TTGAUGCUCUGGAGGGCCCCG	159
28432	HCV-Luc:183U21 TT siRNA Inv	TTGCCACAUGAGUGGCCAAGG	160
28433	HCV-Luc:294U21 TT siRNA Inv	TTGGGAUAGUCCGUCAUGGUG	161
28434	HCV-Luc:293U21 TT siRNA Inv	TTGGAUAGUCCGUCAUGGUGU	162
28435	HCV-Luc:292U21 TT siRNA Inv	TTGAUAGUCCGUCAUGGUGUU	163
28436	HCV-Luc:343L21 TT siRNA (325C) Inv	TTGGGGCCCUCCAGAGCAUCU	164
28437	HCV-Luc:180L21 TT siRNA (182C) Inv	TTGCCUUGGCCACUCAUGUGG	165
28438	HCV-Luc:342L21 TT siRNA (324C) Inv	TTCGGGGCCCUCCAGAGCAUC	166
28439	HCV-Luc:181L21 TT siRNA (183C) Inv	TTCCUUGGCCACUCAUGUGGC	167
28440	HCV-Luc:312L21 TT siRNA (294C) Inv	TTCACCAUGACGGACUAUCCC	168
28441	HCV-Luc:311L21 TT siRNA (293C) Inv	TTACACCAUGACGGACUAUCC	169
28442	HCV-Luc:310L21 TT siRNA (292C) Inv	TTAACACCAUGACGGACUAUC	170
28458	RPI Inverted GL2 Str1 5 5' P=S + TsT	$A_sG_sC_sU_sU_sCAUAAGGCGCAUGC\,T_sT$	171
28459	RPI Inverted GL2 Str2 5 5' P=S + TsT	$G_sC_sA_sU_sG_sCGCCUUAUGAAGCU\,T_sT$	172
28460	RPI GL2 Str1 5 5' P=S + TsT	$C_sG_sU_sA_sC_sGCGGAAUACUUCGA\,T_sT$	173
28461	RPI GL2 Str2 5 5' P=S + TsT	$U_sC_sG_sA_sA_sGUAUUCCGCGUACG\,T_sT$	174
29511	RPI GL2 Str2 + RPI GL2 Str1 (tandem synth. w/ idB on 3' of Str 1)	CGUACGCGGAAUACUUCGATTBUCGAAGUAUUCC GCGUACG TT	175
29543	HBV:248U21 siRNA pos	GUCUAGACUCGUGGUGGACTT	176
29544	HBV:414U21 siRNA pos	CCUGCUGCUAUGCCUCAUCTT	177

29545	HBV:1887U21 siRNA pos	CAAGCCUCCAAGCUGUGCCTT	178
29546	HBV:1877U21 siRNA pos	AGCUGUGCCUUGGGUGGCUTT	179
29547	HBV:228L21 siRNA neg (248C)	GUCCACCACGAGUCUAGACTT	180
29548	HBV:394L21 siRNA neg (414C)	GAUGAGGCAUAGCAGCAGGTT	181
29549	HBV:1847L21 siRNA neg (1867C)	GGCACAGCUUGGAGGCUUGTT	182
29550	HBV:1857L21 siRNA neg (1877C)	AGCCACCCAAGGCACAGCUTT	183
29551	HBV:248U21 siRNA pos Inv	CAGGUGGUGCUCAGAUCUGTT	184
29552	HBV:414U21 siRNA pos Inv	CUACUCCGUAUCGUCGUCCTT	185
29553	HBV:1887U21 siRNA pos Inv	CCGUGUCGAACCUCCGAACTT	186
29554	HBV:1877U21 siRNA pos Inv	UCGGUGGGUUCCGUGUCGATT	187
29555	HBV:228L21 siRNA neg (248C) Inv	CAGAUCUGAGCACCACCUGTT	188
29556	HBV:394L21 siRNA neg (414C) Inv	GGACGACGAUACGGAGUAGTT	189
29557	HBV:1847L21 siRNA neg (1867C) Inv	GUUCCGAAGGUUCGACACGTT	190
29558	HBV:1857L21 siRNA neg (1877C) Inv	UCGACACGGAACCCACCGATT	191
29573	HCV-Luc:182U21 siRNA	CGGAACCGGUGAGUACACCGG	192
29574	HCV-Luc:183U21 siRNA	GGAACCGGUGAGUACACCGGA	193
29575	HCV-Luc:292U21 siRNA	UUGUGGUACUGCCUGAUAGGG	194
29576	HCV-Luc:293U21 siRNA	UGUGGUACUGCCUGAUAGGGU	195
29577	HCV-Luc:294U21 siRNA	GUGGUACUGCCUGAUAGGGUG	196
29578	HCV-Luc:324U21 siRNA	GCCCCGGGAGGUCUCGUAGAC	197
29579	HCV-Luc:325U21 siRNA	CCCCGGGAGGUCUCGUAGACC	198
29580	HCV-Luc:182L21 siRNA (182C)	GGUGUACUCACCGGUUCCGCA	199
29581	HCV-Luc:183L21 siRNA (183C)	CGGUGUACUCACCGGUUCCGC	200
29582	HCV-Luc:312L21 siRNA (282C)	CUAUCAGGCAGUACCACAAGG	201
29583	HCV-Luc:313L21 siRNA (293C)	CCUAUCAGGCAGUACCACAAG	202
29584	HCV-Luc:314L21 siRNA (294C)	CCCUAUCAGGCAGUACCACAA	203
29585	HCV-Luc:344L21 siRNA (324C)	CUACGAGACCUCCCGGGGCAC	204
29586	HCV-Luc:345L21 siRNA (325C)	UCUACGAGACCUCCCGGGGCA	205
29587	HCV-Luc:182U21 siRNA rev	GGCCACAUGAGUGGCCAAGGC	206
29588	HCV-Luc:183U21 siRNA rev	AGGCCACAUGAGUGGCCAAGG	207
29589	HCV-Luc:292U21 siRNA rev	GGGAUAGUCCGUCAUGGUGUU	208
29590	HCV-Luc:293U21 siRNA rev	UGGGAUAGUCCGUCAUGGUGU	209
29591	HCV-Luc:294U21 siRNA rev	GUGGGAUAGUCCGUCAUGGUG	210
29592	HCV-Luc:324U21 siRNA rev	CAGAUGCUCUGGAGGGCCCCG	211
29593	HCV-Luc:325U21 siRNA rev	CCAGAUGCUCUGGAGGGCCCC	212
29594	HCV-Luc:182L21 siRNA (182C) rev	ACGCCUUGGCCACUCAUGUGG	213
29595	HCV-Luc:183L21 siRNA (183C) rev	CGCCUUGGCCACUCAUGUGGC	214
29596	HCV-Luc:312L21 siRNA (282C) rev	GGAACACCAUGACGGACUAUC	215
29597	HCV-Luc:313L21 siRNA (293C) rev	GAACACCAUGACGGACUAUCC	216
29598	HCV-Luc:314L21 siRNA (294C) rev	AACACCAUGACGGACUAUCCC	217
29599	HCV-Luc:344L21 siRNA (324C) rev	CACGGGGCCCUCCAGAGCAUC	218
29600	HCV-Luc:345L21 siRNA (325C) rev	ACGGGGCCCUCCAGAGCAUCU	219
29601	Luc2:128U21 siRNA	CAGAUGCACAUAUCGAGGUGA	220
29602	Luc3:128U21 siRNA	CAGAUGCACAUAUCGAGGUGG	221
29603	Luc2/3:128U21 TT siRNA	CAGAUGCACAUAUCGAGGUTT	222
29604	Luc2/3:148L21 siRNA (128C)	ACCUCGAUAUGUGCAUCUGUA	223
29605	Luc2/3:148L21 TT siRNA (128C)	ACCUCGAUAUGUGCAUCUGTT	224
29606	Luc2/3:166U21 siRNA	UACUUCGAAAUGUCCGUUCGG	225
29607	Luc2/3:186U21 TT siRNA	UACUUCGAAAUGUCCGUUCTT	226
29608	Luc2:186L21 siRNA (168C)	GAACGGACAUUUCGAAGUAUU	227

29609	Luc3:186L21 siRNA (166C)	GAACGGACAUUUCGAAGUACU	228
29610	Luc2/3:186L21 TT siRNA (166C)	GAACGGACAUUUCGAAGUATT	229
29611	Luc2/3:167U21 siRNA	ACUUCGAAAUGUCCGUUCGGU	230
29612	Luc2/3:167U21 TT siRNA	ACUUCGAAAUGUCCGUUCGTT	231
29613	Luc2:187L21 siRNA (167C)	CGAACGGACAUUUCGAAGUAU	232
29614	Luc3:187L21 siRNA (167C)	CGAACGGACAUUUCGAAGUAC	233
29615	Luc2/3:187L21 TT siRNA (167C)	CGAACGGACAUUUCGAAGUTT	234
29616	Luc2/3:652U21 siRNA	AGAUUCUCGCAUGCCAGAGAU	235
29617	Luc2/3:652U21 TT siRNA	AGAUUCUCGCAUGCCAGAGTT	236
29618	Luc2:672L21 siRNA (652C)	CUCUGGCAUGCGAGAAUCUGA	237
29619	Luc3:672L21 siRNA (652C)	CUCUGGCAUGCGAGAAUCUCA	238
29620	Luc2/3:672L21 TT siRNA (652C)	CUCUGGCAUGCGAGAAUCUTT	239
29621	Luc2/3:653U21 siRNA	GAUUCUCGCAUGCCAGAGAUC	240
29622	Luc2/3:653U21 TT siRNA	GAUUCUCGCAUGCCAGAGATT	241
29623	Luc2:673L21 siRNA (653C)	UCUCUGGCAUGCGAGAAUCUG	242
29624	Luc3:673L21 siRNA (653C)	UCUCUGGCAUGCGAGAAUCUC	243
29625	Luc2/3:673L21 TT siRNA (653C)	UCUCUGGCAUGCGAGAAUCTT	244
29626	Luc2/3:880U21 siRNA	UUCUUCGCCAAAAGCACUCUG	245
29627	Luc2/3:880U21 TT siRNA	UUCUUCGCCAAAAGCACUCTT	246
29628	Luc2:900L21 siRNA (880C)	GAGUGCUUUUGGCGAAGAAUG	247
29629	Luc3:900L21 siRNA (880C)	GAGUGCUUUUGGCGAAGAAGG	248
29630	Luc2/3:900L21 TT siRNA (880C)	GAGUGCUUUUGGCGAAGAATT	249
29631	Luc2/3:1012U21 siRNA	CAAGGAUAUGGGCUCACUGAG	250
29632	Luc2/3:1012U21 TT siRNA	CAAGGAUAUGGGCUCACUGTT	251
29633	Luc2:1032L21 siRNA (1012C)	CAGUGAGCCCAUAUCCUUGUC	252
29634	Luc3:1032L21 siRNA (1012C)	CAGUGAGCCCAUAUCCUUGCC	253
29635	Luc2/3:1032L21 TT siRNA (1012C)	CAGUGAGCCCAUAUCCUUGTT	254
29636	Luc2:1139U21 siRNA	AAACGCUGGGCGUUAAUCAGA	255
29637	Luc3:1139U21 siRNA	AAACGCUGGGCGUUAAUCAAA	256
29638	Luc2/3:1139U21 TT siRNA	AAACGCUGGGCGUUAAUCATT	257
29639	Luc2/3:1159L21 siRNA (1139C)	UGAUUAACGCCCAGCGUUUUC	258
29640	Luc2/3:1159L21 TT siRNA (1139C)	UGAUUAACGCCCAGCGUUUTT	259
29641	Luc2:1283U21 siRNA	AAGACGAACACUUCUUCAUAG	260
29642	Luc3:1283U21 siRNA	AAGACGAACACUUCUUCAUCG	261
29643	Luc2/3:1283U21 TT siRNA	AAGACGAACACUUCUUCAUTT	262
29644	Luc2/3:1303L21 siRNA (1283C)	AUGAAGAAGUGUUCGUCUUCG	263
29645	Luc2/3:1303L21 TT siRNA (1283C)	AUGAAGAAGUGUUCGUCUUTT	264
29646	Luc2:1487U21 siRNA	AAGAGAUCGUGGAUUACGUGG	265
29647	Luc3:1487U21 siRNA	AAGAGAUCGUGGAUUACGUCG	266
29648	Luc2/3:1487U21 TT siRNA	AAGAGAUCGUGGAUUACGUTT	267
29649	Luc2/3:1507L21 siRNA (1487C)	ACGUAAUCCACGAUCUCUUUU	268
29650	Luc2/3:1507L21 TT siRNA (1487C)	ACGUAAUCCACGAUCUCUUTT	269
29651	Luc2:1622U21 siRNA	AGGCCAAGAAGGGCGGAAAGU	270
29652	Luc3:1622U21 siRNA	AGGCCAAGAAGGGCGGAAAGA	271
29653	Luc2/3:1622U21 TT siRNA	AGGCCAAGAAGGGCGGAAATT	272
29654	Luc2/3:1642L21 siRNA (1622C)	UUUCCGCCCUUCUUGGCCUUU	273
29655	Luc2/3:1642L21 TT siRNA (1622C)	UUUCCGCCCUUCUUGGCCUTT	274
29656	Luc2:1623U21 siRNA	GGCCAAGAAGGGCGGAAAGUC	275
29657	Luc3:1623U21 siRNA	GGCCAAGAAGGGCGGAAAGAU	276
29658	Luc2/3:1623U21 TT siRNA	GGCCAAGAAGGGCGGAAAGTT	277
29659	Luc2/3:1643L21 siRNA (1623C)	CUUUCCGCCCUUCUUGGCCUU	278

MBHB 02-128B
(900/027)

29660	Luc2/3:1643L21 TT siRNA (1623C)	CUUUCCGCCCUUCUUGGCCTT	279
29663	RPI GL2 Str2, all pyrimidines+ 5BrdUT = PS	$U_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sGU_sT$	280
29664	RPI GL2 Str1 all pyrimidines + 5-BrdUT = PS	$C_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGAU_sT$	281
29665	RPI GL2 Str1 5 5' +5-BrdUT = P=S	$C_sG_sU_sA_sC_sGCGGAAUACUUCGA\ U_sT$	282
29666	RPI GL2 Str2 5' 5 +5BrdUT= P=S	$U_sC_sG_sA_sA_sGUAUUCCGCGUACG\ U_sT$	283
29667	RPI GL2 Str1 all pyrimidines + TT = PS+3'InvAba	$C_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGAT_sTB$	284
29668	RPI GL2 Str1 all pyrimidines = PS+3' and 5' InvAba	$BC_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGAT_sTB$	285
29669	RPI GL2 Str1 all pyrimidines + TT = PS+ 5' InvAba	$BC_sGU_sAC_sGC_sGGAAU_sAC_sU_sU_sC_sGAT_sT$	286
29670	RPI GL2 Str2, all pyrimidines +TT = PS + 3'inverted abasic	$U_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sGT_sTB$	287
29671	RPI GL2 Str2, all pyrimidines +TT = PS + 3'and 5' inverted abasic	$BU_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sGT_sTB$	288
29672	RPI GL2 Str2, all pyrimidines +TT = PS + 5' inverted abasic	$BU_sC_sGAAGU_sAU_sU_sC_sC_sGC_sGU_sAC_sGT_sT$	289
29678	RPI GL2 Str1 + RPI GL2 Str2 (tandem synth. w/ IdB on 3' of Str 2)	UCGAAGUAUUCCGCGUACG TTBCGUACGCGGAAUACUUCGATT	290
29681	RPI GL2 Str1 5'ligation fragment 5-5'-P=S	$C_sG_sU_sA_sC_sG$	291
29682	RPI GL2 Str1 3'-ligation fragment 5-5'-P=S	$CGGAAUACUUCGAT_sT$	292
29683	RPI GL2 Str2 5' ligation fragment 5-5'-P=S	$U_sC_sG_sA_sA_sGUA$	293
29684	RPI GL2 Str2 3' ligation fragment 5-5'-P=S	$UUCCGCGUACGT_sT$	294
29685	RPI GL2 Str2 5' ligation fragment all-P=S	$U_sC_sG_sA_sA_sG_sU_sA$	295
29688	RPI GL2 Str2 3' ligation fragment all-P=S	$U_sU_sC_sC_sG_sC_sG_sU_sA_sC_sG_sT_sT$	296
29694	FLT1:349U21 siRNA stab1	$C_sU_sG_sA_sG_sUUUAAAAGGCACCCT_sT$	297
29695	FLT1:2340U21 siRNA stab1	$C_sA_sA_sC_sC_sACAAAAUACAACAAT_sT$	298
29696	FLT1:3912U21 siRNA stab1	$C_sC_sU_sG_sG_sAAAGAAUCAAAACCT_sT$	299
29697	FLT1:2949U21 siRNA stab1	$G_sC_sA_sA_sG_sGAGGGCCUCUGAUGT_sT$	300
29698	FLT1:369L21 siRNA (349C) stab1	$G_sG_sG_sU_sG_sCCUUUUAAACUCAGT_sT$	301
29699	FLT1:2360L21 siRNA (2340C) stab1	$U_sU_sG_sU_sU_sGUAUUUUGUGGUUGT_sT$	302
29700	FLT1:3932L21 siRNA (3912C) stab1	$G_sG_sU_sU_sU_sUGAUUCUUUCCAGGT_sT$	303
29701	FLT1:2969L21 siRNA (2949C) stab1	$C_sA_sU_sC_sA_sGAGGCCCUCCUUGCT_sT$	304
29706	FLT1:369L21 siRNA (349C) stab2	$G_sG_sG_sU_sG_sC_sC_sU_sU_sU_sU_sA_sA_sA_sC_sU_sC_sA_sG_sT_sT$	305
29707	FLT1:2360L21 siRNA (2340C) stab2	$U_sU_sG_sU_sU_sG_sU_sA_sU_sU_sU_sU_sG_sU_sG_sG_sU_sU_sG_sT_sT$	306
29708	FLT1:3932L21 siRNA (3912C) stab2	$G_sG_sU_sU_sU_sU_sG_sA_sU_sU_sC_sU_sU_sU_sC_sC_sA_sG_sG_sT_sT$	307
29709	FLT1:2969L21 siRNA (2949C) stab2	$C_sA_sU_sC_sA_sG_sA_sG_sG_sC_sC_sC_sU_sC_sC_sU_sU_sG_sC_sT_sT$	308
28030	RPI GL2 Str1	gpcauuggccaacguacgcggaauacuucgauuggguuaggga	309
28242	RPI GL2 Str1 2'-OMe	cguacgcggaauacuucgauu	310

28243	RPI GL2 Str1 14 5' 2'-O-Me	cguacgcggaauacUUCGATT	311
28244	RPI GL2 Str1 10 5' 2'-O-Me	cguacgcggaAUACUUCGATT	312
28245	RPI GL2 Str1 5 5' 2'-O-Me	cguacGCGGAAUACUUCGATT	313
28246	RPI GL2 Str2 all 2'-O-me	ucggaaguauucccgguacguu	314
28247	RPI GL2 Str2 all ribo pyrimidines = 2'-Ome	ucGAAGuAuuccGcGuAcGuu	315
28248	RPI GL2 Str2 5' 14 2'-O-Me	ucggaaguauucccpGUACGTT	316
28249	RPI GL2 Str2 5' 10 2'-O-Me	ucggaaguauuCCCGCGUACGTT	317
28250	RPI GL2 Str2 5' 2'-O-Me	ucggaaGUAUUCCCGCGUACGTT	318
28251	RPI GL2 Str1 all pyrimidines 2'-O-Me except 3'-TT	cGuAcGcGGAAuAcuucGATT	319
28252	RPI GL2 Str1 all pyrimidines = 2'-OMe	cGuAcGcGGAAuAcuucGAuu	320
28253	RPI GL2 Str1+ TT =P=S	CGUACGCGGAAUACUUCGAT$_S$T	321
28281	RPI GL2 Str2 all ribo pyrimidines = 2'-O-me,except 3'-TT	ucGAAGuAuuccGcGuAcGTT	322
28257	RPI GL2 Str1+ 3' univ. base 2	CGUACGCGGAAUACUUCGAXX	323
28258	RPI GL2 Str1+ 3' univ base 1	CGUACGCGGAAUACUUCGAZZ	324
28259	RPI GL2 Str2, + 3' univ. base 2	UCGAAGUAUUCCGCGUACGXX	325
28260	RPI GL2 Str2, + 3' univ. base 1	UCGAAGUAUUCCGCGUACGZZ	326
28014	RPI GL2 Str1 5'ligation fragment P=Scapped Y-2'F	$_S$cG$_S$u$_S$A$_S$cG	327
28015	RPI GL2 Str1 3' ligation fragment P=Scapped Y-2'F	cGGAAuAcuuc$_S$G$_S$A$_S$T$_S$T	328
28028	RPI GL2 Str1 P=Scapped Y-2'F	$_S$cG$_S$u$_S$A$_S$cGcGGAAuAcuuc$_S$G$_S$A$_S$T$_S$T	329
28016	RPI GL2 Str2 5' ligation fragment P=Scapped Y-2'F	u$_S$c$_S$G$_S$A$_S$AGuA	330
28017	RPI GL2 Str2 3'ligation fragment P=Scapped Y-2'F	uucGCGuA$_S$c$_S$G$_S$T$_S$T	331
28027	RPI GL2 Str2 P=Scapped Y-2'F	u$_S$c$_S$G$_S$A$_S$AGuAuucGCGuA$_S$c$_S$G$_S$T$_S$T	332
28018	RPI GL2 Str1 5'ligation fragment 5'P=S Y-2'F	$_S$cGuAcG	333
28019	RPI GL2 Str1 3' ligation fragment 5'P=S Y-2'F	cGGAAuAcuucGATT	334
28029	RPI GL2 Str1 5'P=S Y-2'F	$_S$cGuAcGcGGAAuAcuucGATT	335
28020	RPI GL2 Str2 5' ligation fragment 5'P=S Y-2'F	$_S$ucGAAGuA	336
28021	RPI GL2 Str2 3'ligation fragment 5'P=S Y-2'F	uuccGCGuAcGTT	337
28029	RPI GL2 Str2 5'P=S Y-2'F	$_S$uGAAGuAuuccGCGuAcGTT	338
28022	RPI Inverted GL2 Str1 P=Scapped Y-2'F	A$_S$G$_S$c$_S$u$_S$ucAuAAGGcGcAu$_S$G$_S$c$_S$T$_S$T	339
28023	RPI Inverted GL2 Str2 P=Scapped Y-2'F	G$_S$c$_S$A$_S$u$_S$GcGccuuAuGAAG$_S$c$_S$u$_S$T$_S$T	340
28024	RPI Inverted GL2 Str1 5'P=S Y-2'F	$_S$AGcuucAuAAGGcGcAuGcTT	341
28025	RPI Inverted GL2 Str2 5'P=S Y-2'F	$_S$GcAuGcGccuuAuGAAGcuTT	342
28455	RPI GL2 Str1 2'-F U C	cGuAcGcGGAAuAcuucGATT	343
28456	RPI GL2 Str2 2'-F U C	ucGAAGuAuuccGcGuAcGTT	344
29702	FLT1:348U21 siRNA stab3	c$_S$u$_S$G$_S$A$_S$GuuuAAAAGGcAc$_S$c$_S$c$_S$T$_S$T	345
29703	FLT1:2340U21 siRNA stab3	c$_S$A$_S$A$_S$c$_S$cAcAAAAuAcAAc$_S$A$_S$A$_S$T$_S$T	346
29704	FLT1:3912U21 siRNA stab3	c$_S$c$_S$u$_S$G$_S$GAAAGAAucAAAA$_S$c$_S$G$_S$T$_S$T	347
29705	FLT1:2946U21 siRNA stab3	G$_S$c$_S$A$_S$A$_S$GGAGGGccucuGA$_S$u$_S$G$_S$T$_S$T	348
28443	RPI GL2 Str1 2'-amino U C	cGuAcGcGGAAuAcuucGATT	349
28444	RPI GL2 Str2 2'-amino U C	ucGAAGuAuuccGcGuAcGTT	350

28445	RPI GL2 Str1 2'-amino U C uT 3'end	cGuAcGcGGAAuAcuucGAuT	351
28446	RPI GL2 Str2 2'-amino U C uT 3'end	ucGAAGuAuuccGcGuAcGuT	352
30051	HCV-Luc:325U21 siRNA 5 5' P=S + 3' univ. base 2 + 5'/3' invAbs	BC$_S$C$_S$C$_S$C$_S$G$_S$GGAGGUCUCGUACGAXXB	353
30052	HCV-Luc:325U21 siRNA rev 5 5' P=S + 3' univ. base 2 + 5'/3' invAbs	BA$_S$G$_S$A$_S$U$_S$G$_S$CUCUGGAGGGCCCCXXB	354
30053	HCV-Luc:345L21 siRNA (325C) 5 5' P=S + 3' univ. base 2 + 3' invAbs	U$_S$C$_S$U$_S$A$_S$C$_S$GAGACCUCCCGGGGXXB	355
30054	HCV-Luc:345L21 siRNA (325C) rev 5 5' P=S + 3' univ. base 2 + 3' invAbs	G$_S$G$_S$G$_S$G$_S$C$_S$CCUCCAGAGCAUCUXXB	356
30055	HCV-Luc:325U21 siRNA all Y P=S + 3' univ. base 2 + 5'/3' invAbs	BC$_S$C$_S$C$_S$C$_S$GGGAGGU$_S$C$_S$U$_S$C$_S$GU$_S$AGAXXB	357
30056	HCV-Luc:325U21 siRNA rev all Y P=S + 3' univ. base 2 + 5'/3' invAbs	BAGAU$_S$GC$_S$U$_S$C$_S$U$_S$GGAGGGC$_S$C$_S$C$_S$C$_S$XXB	358
30057	HCV-Luc:345L21 siRNA (325C) all Y P=S + 3' univ. base 2 + 3' invAbs	U$_S$C$_S$U$_S$AC$_S$GAGAC$_S$C$_S$U$_S$C$_S$C$_S$C$_S$GGGGXXB	359
30058	HCV-Luc:345L21 siRNA (325C) rev all Y P=S + 3' univ. base 2 + 3' invAbs	GGGGC$_S$C$_S$C$_S$U$_S$C$_S$C$_S$AGAGC$_S$AU$_S$C$_S$U$_S$XXB	360
30059	HCV-Luc:325U21 siRNA 4/3 P=S ends + all Y-2'F + 3' univ. base 2 + 5'/3' invAbs	Bc$_S$c$_S$c$_S$c$_S$GGGAGGucuuGuA$_S$G$_S$A$_S$XXB	361
30060	HCV-Luc:325U21 siRNA rev 4/3 P=S ends + all Y-2'F + 3' univ. base 2 + 5'/3' invAbs	BA$_S$G$_S$A$_S$u$_S$GcucuGGAGGGcc$_S$c$_S$c$_S$XXB	362
30170	HCV-Luc:325U21 siRNA all Y-2'F + 3' univ. base 2 + 5'/3' invAbs	B ccccGGGAGGucuuGuAGAXX B	363
30171	HCV-Luc:325U21 siRNA rev all Y-2'F + 3' univ. base 2 + 5'/3' invAbs	B AGAuGcucuGGAGGGccccXX B	364
30172	HCV-Luc:345L21 siRNA (325C) all Y P=S + 3' univ. base 2 + 5'/3' invAbs	B U$_S$C$_S$U$_S$AC$_S$GAGAC$_S$C$_S$U$_S$C$_S$C$_S$C$_S$GGGGXX B	365
30173	HCV-Luc:345L21 siRNA (325C) all Y-2'F	ucuAcGAGAccucccGGGG	366
30174	HCV-Luc:345L21 siRNA (325C) rev all Y-2'F	GGGGcccucccAGAGcAucu	367
30175	HCV-Luc:345L21 siRNA (325C) all Y-2'F + 3' univ. base 2	ucuAcGAGAccucccGGGGXX	368
30176	HCV-Luc:345L21 siRNA (325C) rev all Y-2'F + 3' univ. base 2	GGGGcccucccAGAGcAucuXX	369
30177	HCV-Luc:345L21 siRNA (325C) all Y-2'F + 3' univ. base 2 + 5'/3' IB	B ucuAcGAGAccucccGGGGXX B	370
30178	HCV-Luc:325U21 siRNA all Y P=S + 3' univ. base 2 + 3' invAbs	C$_S$C$_S$C$_S$C$_S$GGGAGGU$_S$C$_S$U$_S$C$_S$GU$_S$AGAXX B	371

UPPER CASE = ribonucleotide
Lower case = 2'-O-methyl nucleotide
Underline = 2'-deoxy-2'-amino nucleotide
Italic = 2'-deoxy-2'-fluro nucleotide
T = thymidine
B = inverted deoxyabasic succinate linker
B = inverted deoxyabasic
X = universal base (5-nitroindole)
Z = universal base (3-nitropyrrole)
S = phosphorothioate internucleotide linkage
U = 5-bromodeoxyuridine

Table II

A. 2.5 µmol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	6.5	163 µL	45 sec	2.5 min	7.5 min
S-Ethyl Tetrazole	23.8	238 µL	45 sec	2.5 min	7.5 min
Acetic Anhydride	100	233 µL	5 sec	5 sec	5 sec
N-Methyl Imidazole	186	233 µL	5 sec	5 sec	5 sec
TCA	176	2.3 mL	21 sec	21 sec	21 sec
Iodine	11.2	1.7 mL	45 sec	45 sec	45 sec
Beaucage	12.9	645 µL	100 sec	300 sec	300 sec
Acetonitrile	NA	6.67 mL	NA	NA	NA

B. 0.2 µmol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	15	31 µL	45 sec	233 sec	465 sec
S-Ethyl Tetrazole	38.7	31 µL	45 sec	233 min	465 sec
Acetic Anhydride	655	124 µL	5 sec	5 sec	5 sec
N-Methyl Imidazole	1245	124 µL	5 sec	5 sec	5 sec
TCA	700	732 µL	10 sec	10 sec	10 sec
Iodine	20.6	244 µL	15 sec	15 sec	15 sec
Beaucage	7.7	232 µL	100 sec	300 sec	300 sec
Acetonitrile	NA	2.64 mL	NA	NA	NA

C. 0.2 µmol Synthesis Cycle 96 well Instrument

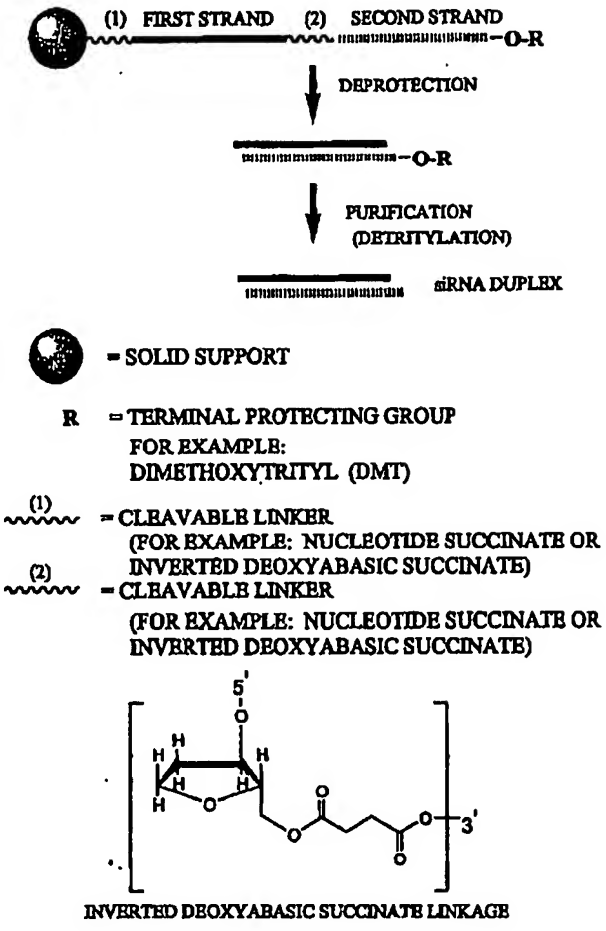
Reagent	Equivalents:DNA/ 2'-O-methyl/Ribo	Amount: DNA/2'-O-methyl/Ribo	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time* Ribo
Phosphoramidites	22/33/66	40/60/120 µL	60 sec	180 sec	360sec
S-Ethyl Tetrazole	70/105/210	40/60/120 µL	60 sec	180 min	360 sec
Acetic Anhydride	265/265/265	50/50/50 µL	10 sec	10 sec	10 sec
N-Methyl Imidazole	502/502/502	50/50/50 µL	10 sec	10 sec	10 sec
TCA	238/475/475	250/500/500 µL	15 sec	15 sec	15 sec
Iodine	8.8/8.8/8.8	80/80/80 µL	30 sec	30 sec	30 sec
Beaucage	34/51/51	80/120/120	100 sec	200 sec	200 sec
Acetonitrile	NA	1150/1150/1150 µL	NA	NA	NA

- Wait time does not include contact time during delivery.
- Tandem synthesis utilizes double coupling of linker molecule

ABSTRACT OF THE DISCLOSURE

The present invention concerns methods and reagents useful in modulating gene expression in a variety of applications, including use in therapeutic, diagnostic, agricultural, target validation, and genomic discovery applications. Specifically, the invention relates to synthetic chemically modified small interfering RNA (siRNA) molecules capable of mediating RNA interference (RNAi).

Figure 1



INVERTED DEOXYABASIC SUCCINATE LINKAGE

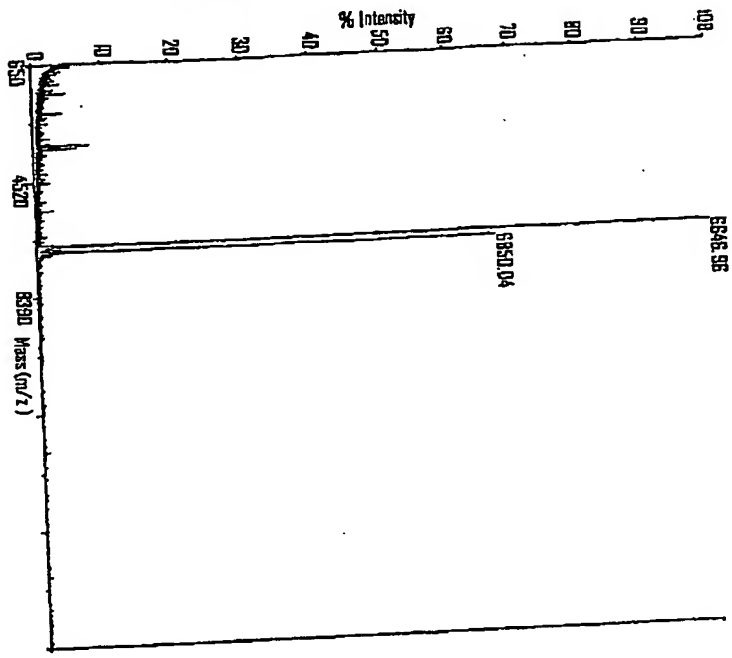


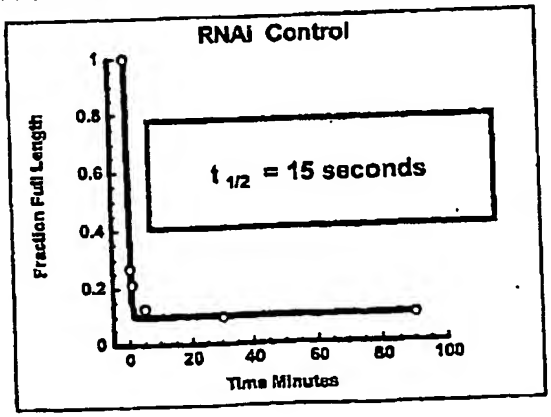
Figure 2

Figure 3

CGUACGCGGAAUACUUCGATT (SEQ ID NO: 372) T ½ = 15 seconds (RNAi Control)
TTGCAUGCGCCUUAUGAAGCU (SEQ ID NO: 373)

$C_sG_sU_sA_sCGCGGAAUACUUC_sG_sA_sT_sT$ (SEQ ID NO: 374) T ½ = 24 hours
$T_sT_sG_sC_sAUGCGCCUUAUGA_sA_sG_sC_sU$ (SEQ ID NO: 375) = 329 in 01

$_sCGUACGCGGAAUACUUCGATT$ (SEQ ID NO: 376) T ½ = 72 hours = 332 in 01
$TTGCAUGCGCCUUAUGAAGCU_s$ (SEQ ID NO: 377)



G = Guanosine
A = Adenosine
U = Uridine
C = Cytidine
T = Thymidine
Lower Case = 2'-deoxy-2'-fluoro
S = phosphorothioate

Figure 4

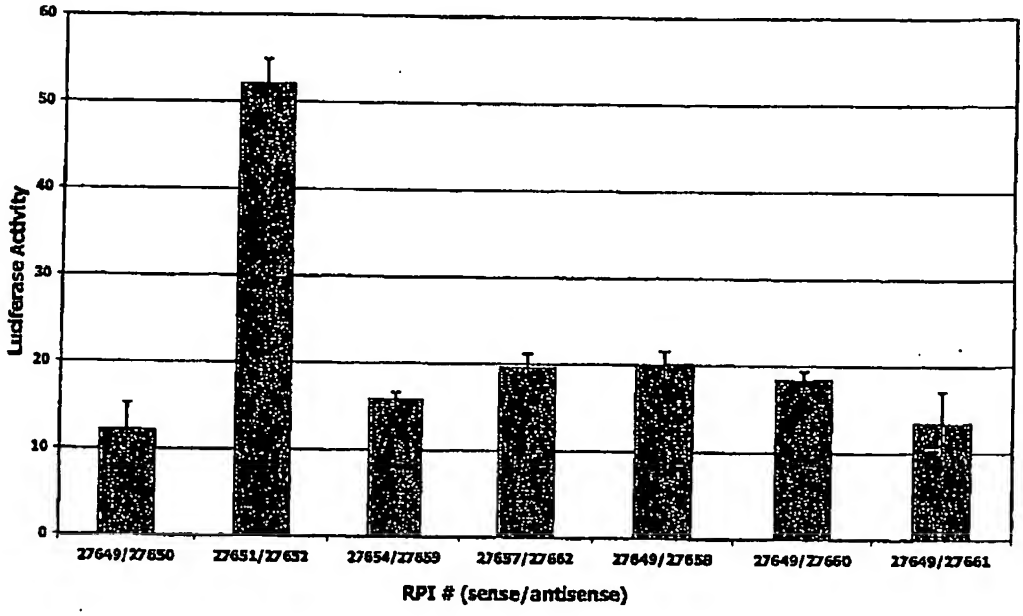


Figure 5

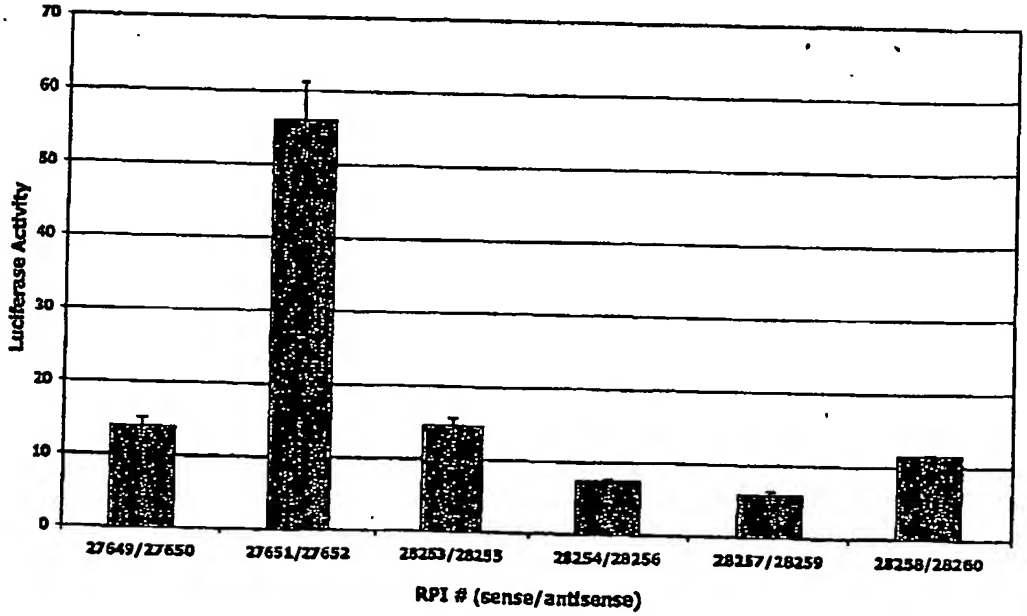


Figure 6

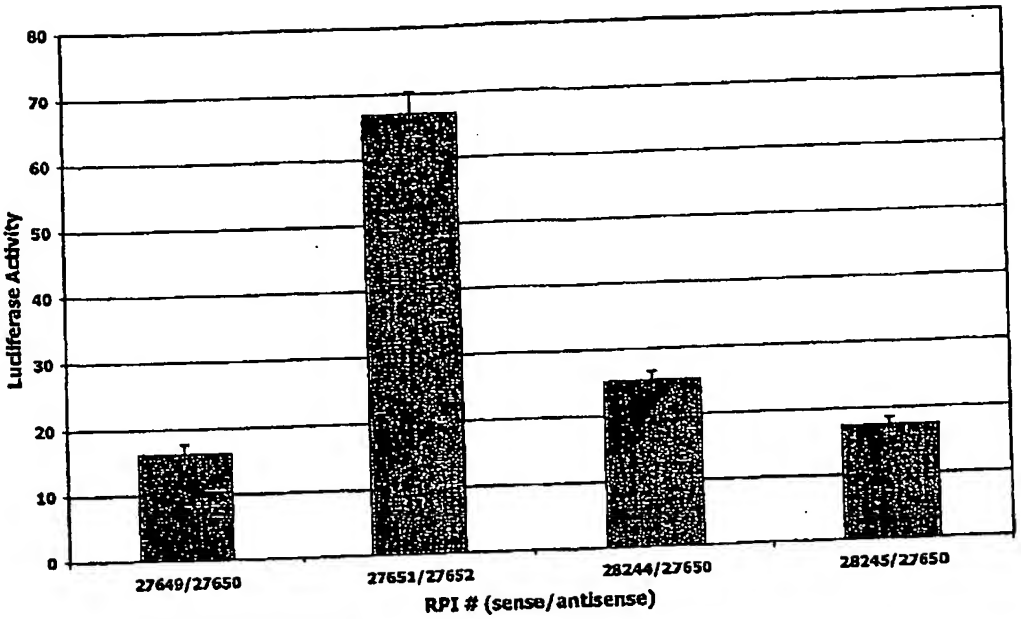


Figure 7

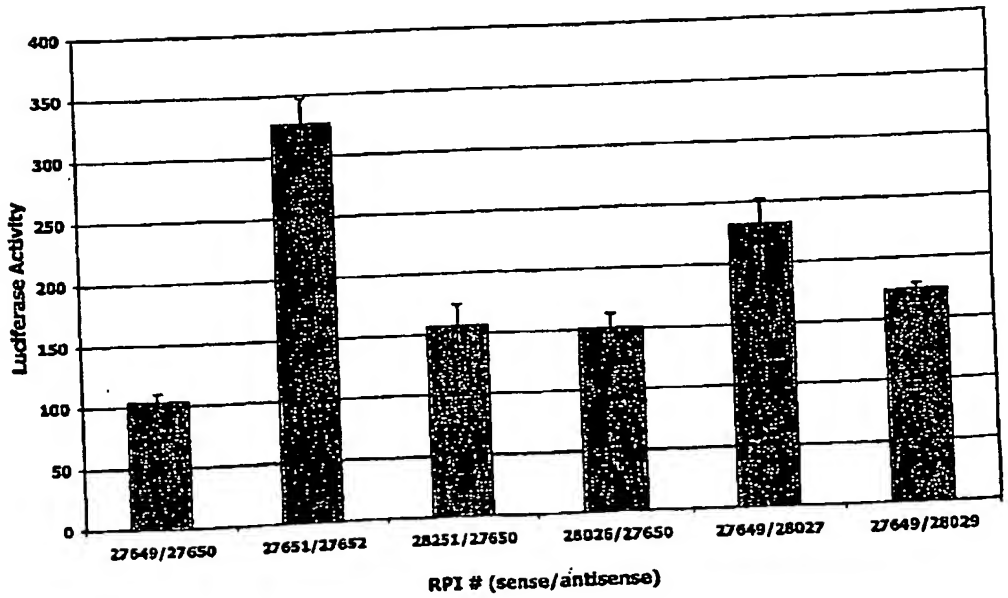


Figure 8

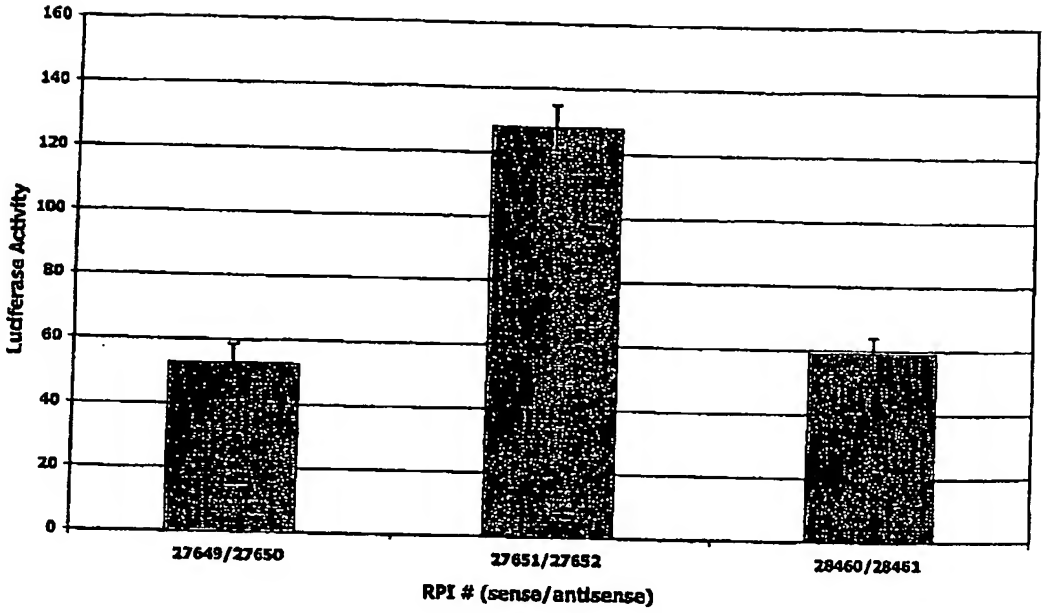
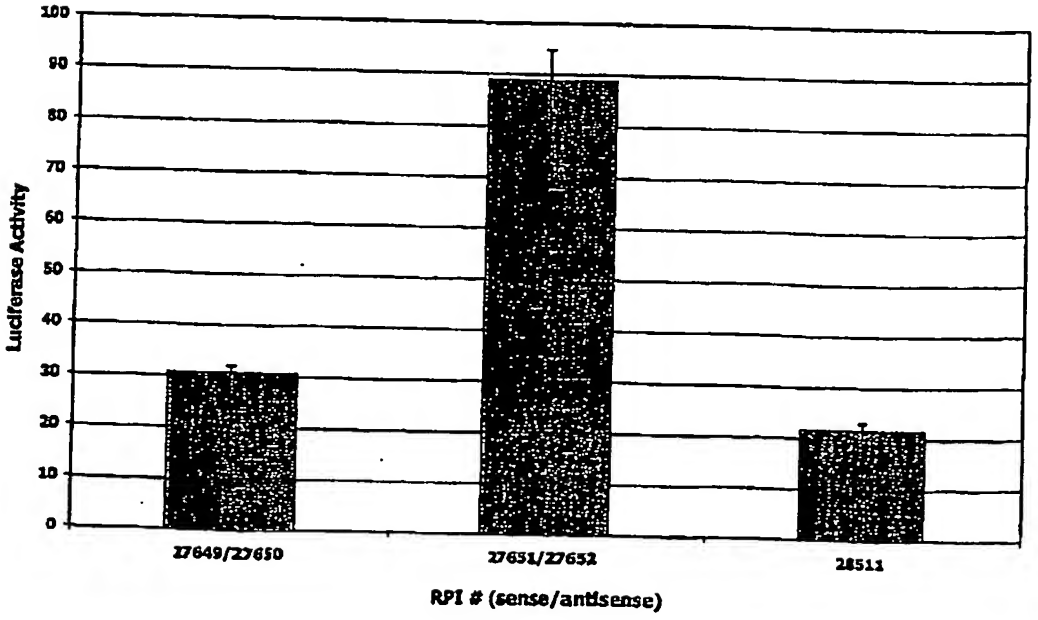
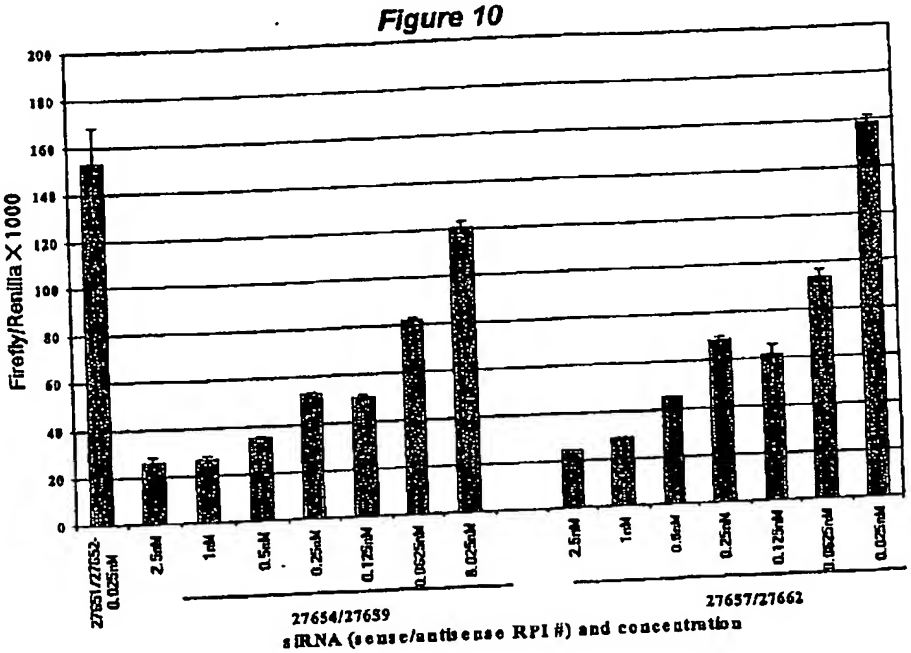


Figure 9



209090*28798E09

Figure 10



209090*28798E09

Figure 11: siRNAs targeting HCV chimera

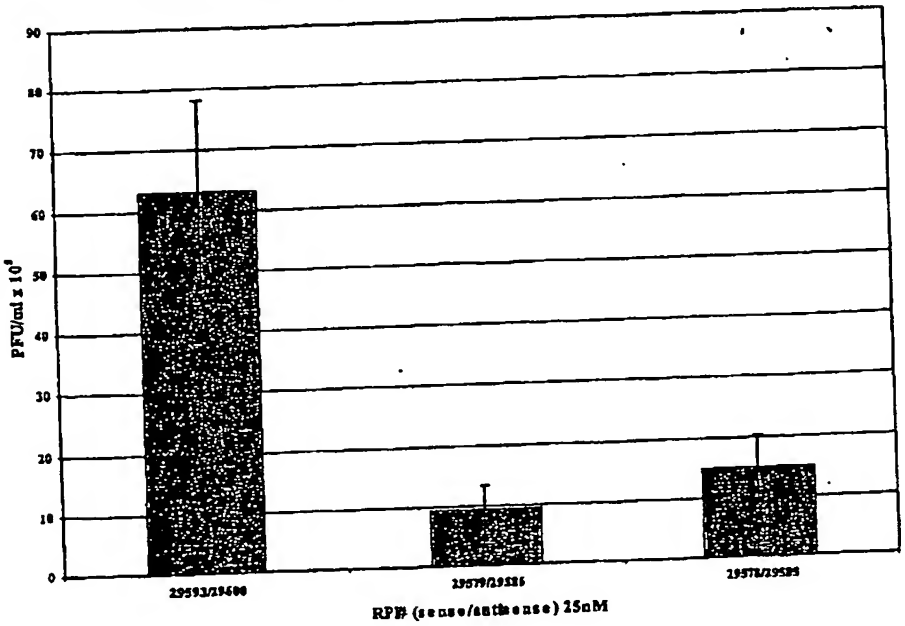


Figure 12: HCV siRNA dose response

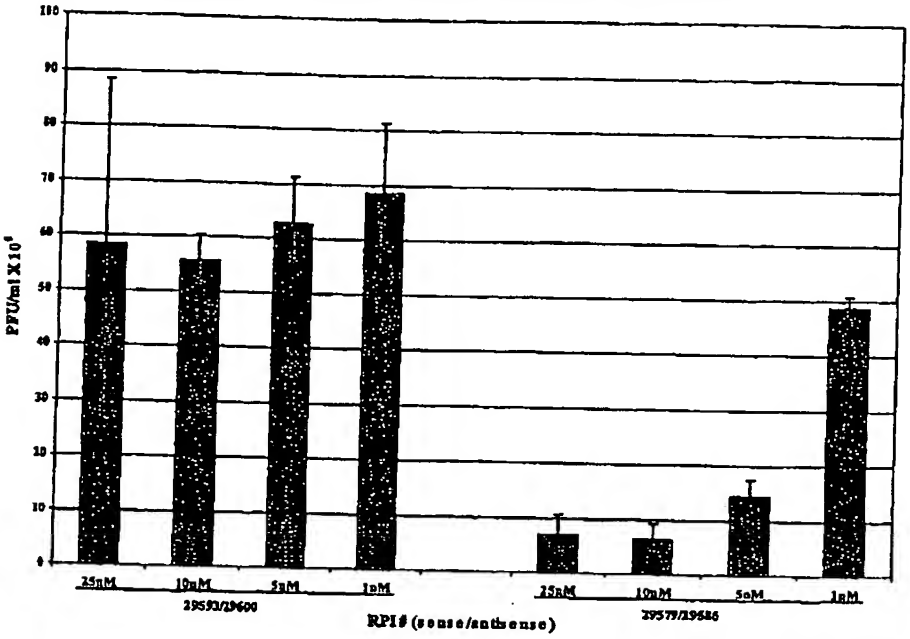


Figure 13: Chemically Modified siRNA targeting HCV chimera

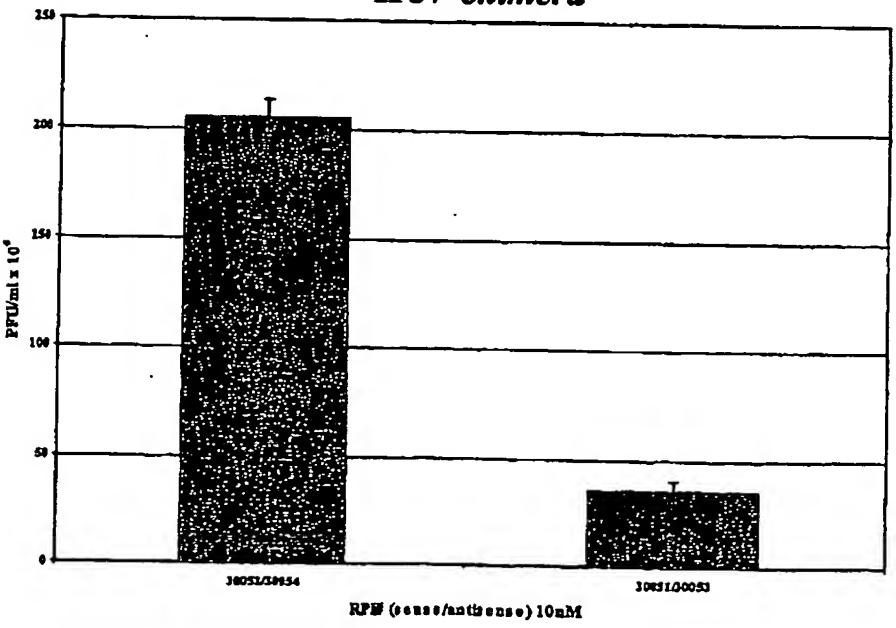


Figure 14: Chemically Modified siRNA targeting HCV chimera